DOCUMENT RESUME

ED 177 416 CG 013 844

AUTHOR	Waller, John D.; And Others
TITLE	Developing Useful Evaluation Capability: Lessons From the Model Evaluation Program.
INSTITUTION	Urban Inst., Washington, D.C.
SPONS AGENCY	National Inst. of Law Enforcement and Criminal Justice (Dept. of Justice/LEAA), Washington, D.C.
PUB DATE	Jun 79
GRANT	77-N1-99-0022
NOTE	161p.; Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (027-000-00798-8)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	*Community Agencies (Public); *Demonstration Projects; Efficiency; *Evaluation Methods; *Government (Administrative Body); *Grants; *Organizational Effectiveness; Program Effectiveness

ABSTRACT

The assessment of 12 model evaluation systems provides insight and guidance into their development for government managers and evaluators. The eight individual completed grants are documented in a series of case studies, while the synthesis of all the project experiences and results are summarized. There are things that evaluation systems can do and cannot do. Their success was gauged by: (1) degree to which helping agencies achieve their objectives; (2) identifying and meeting a real demand for evaluation information; and (3) continuing after grant expiration. No site met the first criteria, but all save one site produced evaluation information, and three of the eight tested systems were continued after grant expiration; however their long-term fate is uncertain. During the building of an evaluation capability phase, expectations and strategies often needed changing. Site characteristics, also, varied considerably. Many problems may be expected in implementation and operation of evaluation systems, including the fact that there are neither standard problems nor easy answers. (LS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED177416

Developing Useful Evaluation Capability: Lessons From the Model Evaluation Program

by
John D. Waller
John W. Scanlon
Dona M. Kemp
Paul G. Nalley

June 1979

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CG013844

U. S. Department of Justice
Law Enforcement Assistance Administration
National Institute of Law Enforcement and Criminal Justice

Law Enforcement Assistance Administration
Henry S. Dogin, *Administrator*

National Institute of Law Enforcement
and Criminal Justice
Harry M. Bratt, *Acting Director*

This project was supported by Grant Number 77-NI-99-0022, awarded to The Urban Institute by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U. S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Research on this project was completed in April 1978. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U. S. Department of Justice.

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Stock Number 027-000-00796-8

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS viii

PREFACE ix

I. LESSONS DRAWN FROM THE MODEL EVALUATION PROGRAM 1

- A. What Evaluation Systems Can and Cannot Do 2
  1. Evaluation Systems Cannot Be Expected To Have a Measurable Impact on the Performance of State Planning Agencies and Regional Planning Units, since these Agencies Do Not Have Objectives Regarding their Primary Mission that Are both Verifiable and Realistic. 2
  2. Evaluation Systems Have an even Chance of Providing Information that Will Be Used by either Planning Agencies or Criminal Justice Agencies. 4
  3. Evaluation Systems Are not Likely To Be Useful to the Planning Agency Staff unless the Agency Is Being Directed toward some Mission-Related Objective. 5
  4. Evaluation Systems Have an Excellent Chance of Providing Useful Information to State and Local Criminal Justice Agencies if They Chose to Serve this Audience. 5
  5. An Evaluation System Can Increase its Chances of Producing Useful Information by: Serving Users and Information Needs that Already Exist, and Allowing the Users to Own the Study. 6
  6. Planning Agency Evaluation Systems Carrying Out Intensive Evaluations to Test Program Concepts Will Not Be Successful. 7
- B. What Evaluation Systems Can and Cannot Be Set Up 8
  1. If Funds Are Earmarked for Evaluation, Planning Agencies Are Almost Certain To Set Up an Internal Evaluation System that Produces, or Helps Produce, Information. 9
  2. Development of an Evaluation Capability Does Not Assure Institutionalization of the Capability. 9
  3. Planning Agencies Appear To Be Unsuccessful in Establishing Evaluation Systems in Local Criminal Justice Agencies. 10
  4. Given the Opportunity, Planning Agencies Will Set Up Evaluation Systems Which Differ from Each Other in Product, Cost, and Operations. 10
  5. The Only Characteristic of an Evaluation System Associated with Utility Was the Degree of Involvement of the User in the Evaluation Activity. 11

TABLE OF CONTENTS (Cont'd.)

Page

C. What Operational Problems May Arise in Setting Up and Operating Evaluation Systems 12
  1. Staff Turnover in the Evaluation Unit Will Cause Delays and Disruptions. 13
  2. Civil Service Rules May Cause Delays. 13
  3. Projects May Be Reluctant to Report Data. 13
  4. Project Files May Contain Less Data Than Expected or Required for Planned Evaluation. 14
  5. Computer Processing of Data Can Be a Time Consuming, Error-Plagued Undertaking. 14
  6. Contracting Out Evaluations Has Its Own Set of Problems. 15
  7. The Identified Uses Can Add to the Cost of Evaluation by Having Unclear or Shifting Information Requirements. 16
  8. Diversion of Evaluation Unit Staff to Other Planning Agency Function Is a Danger Signal. 16
D. The MEP Message: User Involvement 17

II. SUCCESS OF THE TESTED EVALUATION SYSTEMS 18

A. Overview of the Model Evaluation Program Outcomes 18
B. Three Success Criteria for the Model Evaluation Programs: Achievement of Agency Objectives, Use and Continuation 22
C. MEP Contribution to the Achievement of Agency Objectives Not Demonstrable 25
D. Half the MEP Sites Were Able To Demonstrate that a Market for Evaluation Was Served 30
  1. Some Limitations of the Use Criterion 31
  2. Types of Market for Evaluation 32
  3. Project Oversight 35
  4. Issue Identification 50
  5. Knowledge Development 70
E. Site Continuation of the Evaluation System Tested 72
  1. Continuation: Michigan, Jacksonville, and Ventura 74
  2. Discontinued 77
F. Factors Which Affected Success 83
  1. Success in Having Tested System Continued 83
  2. Success in Producing Useful Information 86

III. BUILDING AN EVALUATION CAPABILITY 89

A. Planning Agencies Were Able To Establish Evaluation Systems 89
B. Expectations and Strategies Often Had To Be Changed 91
  1. Changing Expectations and Strategies 92
  2. Types of Strategy Followed 97

TABLE OF CONTENTS (Cont'd.)

Page

C. Cost Items Involved in Setting Up and Operating Evaluation Systems 98
D. Evaluation System Characteristics Varied Considerably across MEP Sites 103
  1. Evaluation Production Levels Achieved 104
  2. Product of the Planning Agency's Evaluation Unit 108
  3. Standardization of the Production Process 110
  4. Evaluation Methodologies 111
  5. Quality Control Mechanisms 113
  6. Mechanisms To Involve the User 117
E. Brief Descriptions of the Evaluation Systems 120
  1. A Standardized SPA Project Reporting System Developed by the Evaluation Unit: Illinois 121
  2. A Standard SPA Project Reporting System Developed by Program Staff: Michigan 121
  3. Phased Contracts To Do Intensive Evaluations: Michigan 122
  4. In-House Staff Do Short-Term Evaluation on Demand: Jacksonville 122
  5. In-House Staff Do Project Evaluations Using Standard Framework: Central Oklahoma 123
  6. RPU Research Teams Do Studies under the Direction of Users: Alameda 124
  7. Advice and Assistance on Evaluation to Subgrantees: Columbia 125
  8. Evaluation Technical Assistance to Local Agencies on Request: Ventura 126
F. Conditions Which May Affect Implementation 127

IV. PROBLEMS TO BE EXPECTED IN IMPLEMENTING AND OPERATING EVALUATION SYSTEMS 129

A. Staff Turnover in the Evaluation Unit Will Cause Delays 129
B. Civil Service Rules May Cause Delays 131
C. Projects May Be Reluctant to Report Data 132
D. Project Files May Contain Less Data than Expected or Required for Planned Evaluation 133
E. Computer Processing of Data Can Be a Time Consuming, Error-Plagued Undertaking 134
F. Contracting Out Evaluations Has its Own Set of Problems 136
G. The Identified User Can Add to the Cost of Evaluation by Having Unclear or Shifting Information Requirements 137
H. Diversion of Evaluation Unit Staff to Other Planning Agency Functions Is a Danger Signal 139
I. No Standard Problem and No Easy Answers Demonstrated 140

TABLE OF CONTENTS (Cont'd.)

Page

LIST OF FIGURES

Figure 1: Outcomes of the Model Evaluation Program 20

Figure 2: Production Levels Achieved by MEP Sites 21

Figure 3: Illustration of One Type of Agency Objective Being Tracked over Time 28

Figure 4: Evaluation System Operations, Information Flows, and Resource Flows 106

TABLE OF CONTENTS (Cont'd.)

Page

LIST OF TABLES

Table 1: Type of Use for Which MEP Sites Planned to Develop Informatin and Were Able to Develop Information 36

Table 2: Success in Producing Useful Information 37

Table 3: Continuation of the Tested System 73

Table 4: Budgets and Staffs of Model Evaluation Program Grants 85

Table 5: Sites that Met or Changed Expectations 93

Table 6: Implementation Strategies Followed in Setting Up the Tested System 99

Table 7: Cost Items Incurred in Setting Up and Operating MEP Evaluation Systems 101-102

Table 8: MEP Evaluation System Characteristics 105

Table 9: Production Levels Achieved by MEP Sites 107

Table 10: Methods for Performing Evaluation Operations in the MEP 112

Table 11: Ventura MEP - Frequency of Use of Different Designs (Measure and Comparison) in Eighteen Studies 114

Table 12: Summary of Michigan's Evaluation Design for Youth Service Bureaus 115-116

Table 13: Operational Problems that Evaluation Management Should Be Prepared to Resolve 130

ACKNOWLEDGMENTS

This report is based on the experiences of the 12 Model Evaluation Program (MEP) grantees that developed and implemented their approaches to evaluation. Without their willingness to experiment with new approaches and be subject to evaluation themselves, this report and the lessons contained in it would not have been possible.

The authors wish to recognize the MEP participants for their openness and cooperation in this program. They provided extensive documentation on their projects, and submitted to extensive interviews and surveys. The authors are grateful for the time and effort expended. Unfortunately, the list of persons who provided information over a several-year period is too numerous to give here.

The MEP and this assessment also depended on the efforts and staff of The National Institute of Law Enforcement and Criminal Justice (NILECJ). The initial grant to The Urban Institute was awarded when Gerald Caplan was Director of the National Institute, and was continued by Blair Ewing. The principal NILECJ staff member on the MEP was Dr. Paul Lineberry, who expended many hours with the authors and in overseeing the grantees. He provided many insights into the management and evaluation activities studied, and continually pushed for a synthesis that would provide useful guidance to all government managers and evaluators.

Other National Institute personnel contributing to the MEP and this effort are: Joel Garner; Victoria Jaycox; Richard Rau; and John Spevacek. Their support and cooperation were essential and are appreciated.

Finally, many members of the Program Evaluation Research Program of The Urban Institute contributed to the MEP and this report by performing field work or reviewing products. They include: Richard Schmidt; Joseph Wholey; and Sumner Clarren.

PREFACE

This report presents findings from an assessment of 12 attempts to develop and demonstrate model evaluation systems. Some were successful, some were not, and four are still underway at this time. The experiences encountered provide insights and guidance for those with a responsibility or interest in having evaluation make a difference in government operations.

The 12 model efforts were funded through grants from The National Institute of Law Enforcement and Criminal Justice (NILECJ), and constitute The National Institute's Model Evaluation Program (MEP). The grantees were selected through a competitive process from proposals submitted by State Planning Agencies (SPAs) and Regional Planning Units (RPUs). The SPAs and RPUs are agencies established by the Omnibus Crime Control and Safe Streets Act of 1968 to carry out the planning and funding process authorized by the Act. During recent fiscal years, these agencies participated in the annual expenditure of over six-hundred million dollars in grants for projects at the state and local levels of government. In many cases, the SPAs and RPUs are components of state and regional organizations performing other functions than those authorized by the Safe Streets Act. For example, the RPU is part of a Council of Governments or some other regional organization of local governments.

The SPAs and RPUs have responsibility for developing comprehensive plans for the expenditure of funds awarded to them, and for administering the award and expenditure of grants created with those funds. The specific roles and responsibilities of the SPAs and RPUs vary among states and regions. The professional staffs of SPAs range from four to nearly seventy. Funds are awarded to states on a formula basis, with the larger states receiving the largest allocation.

The Law Enforcement Assistance Administration (LEAA) has overall responsibility for administering the Omnibus Crime Act. The authorizing legislation and LEAA placed monitoring and evaluation requirements on the SPAs, who in turn place requirements on or receive cooperation from the RPUs. Over the years, LEAA, the SPAs and RPUs have emphasized evaluation, and have invested heavily in it. The Model Evaluation Program is but one of the LEAA and National Institute efforts to affect performance in the criminal justice area through evaluation.

In developing proposals for the $2,000,000 set aside for MEP grants, SPAs and RPUs were given extensive freedom. As stated in the Program Announcement, "Any proposal to develop a system for generating information on the results, costs and effectiveness of criminal justice projects, programs and activities is eligible."[1] The program objectives were:

- encourage state and local agencies to generate and use evaluation information; and
- test in what ways the effective use of evaluation information can help state and local agencies achieve their objectives.

Additional criteria and instructions were given in the announcement, but the freedom to propose what the SPAs or RPUs believed most appropriate was maintained and exhibited in the variety of proposals funded.

The 12 agencies receiving MEP grants, the short titles used to reference them in this report, and the status as of January 1978 are:

- State Planning Agencies:
  - (1) The Illinois Law Enforcement Commission (short title used here--Illinois; status--grant completed);
  - (2) The Massachusetts Committee on Criminal Justice (short title used here--Massachusetts; status--grant in progress);
  - (3) The Michigan Office of Criminal Justice Programs (short title used here--Michigan; status--grant completed);
  - (4) The New Hampshire Governor's Commission on Crime and Delinquency (short title used here--New Hampshire; status--grant in progress);
  - (5) The Pennsylvania Governor's Justice Commission (short title used here--Pennsylvania; status--grant completed);
  - (6) The Virginia Division of Justice and Crime Prevention (short title used here--Virginia; status--grant in progress);
  - (7) The Washington Law and Justice Planning Office of the Community Development Office (short title used here--Washington; status--grant in progress).
- Regional Planning Units:
  - (1) The Alameda (California) Regional Criminal Justice Planning Board (short title used here--Alameda; status--grant completed);

---

[1] Model Evaluation Program Announcement, Washington, D.C., Law Enforcement Assistance Administration, Fall 1974.

(2) The Association of Central Oklahoma Governments (short title used here--Central Oklahoma; status--grant completed);

(3) The Columbia, South Carolina Central Midlands Regional Planning Council (short title used here--Columbia; status--grant completed);

(4) The Jacksonville, Florida Office of Criminal Justice Planning (short title used here--Jacksonville; status--grant completed);

(5) The Ventura Criminal Justice Planning Board (short title used here--Ventura; status--grant completed).

The Urban Institute was selected to document and assess the MEP. The work performed in that capacity included assistance and documentation of the individual grantee activities, and synthesizing the experiences and results of all the grantees. The individual documentation of the eight completed grants is contained in a series of case studies delivered to the National Institute. This document reports on the synthesis work, and summarizes the experiences and results of the individual projects. It principally presents information on the eight sites that had completed their grants by January 1978, but reflects experiences of all twelve sites.

Over the years the authors have worked with a broad range of federal, state and local agencies, and drew on those experiences in preparing this document. It is believed that the findings and conclusions based on the 12 MEP sites would not be significantly different if other sites had been included, and that the contents are applicable to most government agencies.

I. LESSONS DRAWN FROM THE MODEL EVALUATION PROGRAM

The Model Evaluation Program (MEP) of the Law Enforcement Assistance Administration (LEAA) was successful in identifying both the potential and the limitations of a diverse set of evaluation systems, and the types of operational problem with which evaluation management must be prepared to resolve.

This document summarizes and synthesizes the experiences of state and local agencies that participated in the MEP. The presentation is aimed at persons who may wish to develop an evaluation capability in a government agency, manage an evaluation system, or are part of such a system.

This chapter summarizes the lessons drawn from the MEP. The lessons are organized in three sections: (1) "what evaluation systems can and cannot do" presents lessons on the utility of evaluation; (2) "what evaluation capabilities can and cannot be set up" presents lessons on the implementation of evaluation systems; and (3) "what operational problems may arise" presents lessons on the problems that evaluation managers should be prepared to resolve. In the discussion of the lessons, the term evaluation system refers to systems set up by the State Planning Agency (SPA) or Regional Planning Unit (RPU) during the MEP.

The lessons presented here are supported and discussed in the remainder of the document. Background information on the Model Evaluation Program and program participants is given in the Preface. Also, brief descriptions of the systems developed by the participants are given in Section

D of Chapter III. Those unfamiliar with the MEP may wish to review this material before proceeding with the lessons learned. The lessons reflect conditions encountered in the MEP, but are given as universal statements, since they coincide with similar conditions found in other government agencies.

A. WHAT EVALUATION SYSTEMS CAN AND CANNOT DO

Three outcomes anticipated for the MEP sites were:

- production of evaluation information for specific uses (which they specified), with
- utilization of the information to
- achieve agency objectives.

Successes were realized on the first two outcomes, but not on the third. The successes and failures of the MEP sites on these outcomes are the bases for the following six lessons on what evaluation systems can accomplish.

1. EVALUATION SYSTEMS CANNOT BE EXPECTED TO HAVE A MEASURABLE IMPACT ON THE PERFORMANCE OF STATE PLANNING AGENCIES AND REGIONAL PLANNING UNITS, SINCE THESE AGENCIES DO NOT HAVE OBJECTIVES REGARDING THEIR PRIMARY MISSION THAT ARE BOTH VERIFIABLE AND REALISTIC. (No MEP site was judged to have such objectives.)

One purpose of the MEP was "to test in what ways the effective use of evaluation information can help state and local (planning) agencies achieve their objectives." For LEAA funded agencies, the authorizing legislation specifies the reduction of crime and the improvement of the criminal justice system as primary missions. Planning agencies involved in the MEP provided statements of their objectives and were interviewed about the relationship between the MEP and agency objectives. These objectives turned out not to be in terms that could be measured and documented. Furthermore, in many cases, it was impossible to determine what the agency was trying to accomplish--other than disburse grant funds and meet federal administrative requirements.

Several of the 12 MEP agencies appeared to be well administered, with a sense of purpose, direction and progress. As with other agencies, the agencies judged to be well administered had objectives which had not been translated into measurable terms-"move local agencies toward adoption of national goals and standards," "increase the number of successful innovative projects continued by local agencies." However, the agencies had staff functions in place which appeared to be trying to bring about these objectives. That is, significant portions of the planning agency administrative budget and management attention seemed plausibly linked to the types of rhetorical objective put forward. The difficulty was that even in these cases, the objectives and progress are not measurable or verifiable. Where agencies appeared to be well administered toward some objective, it always seemed straightforward to translate rhetorical objectives into verifiable measures, even though the agencies had not taken that step.

None of the MEP demonstrations was used to formulate measurable agency objectives and evaluate progress in achieving them. In the absence of such objectives, it is impossible to say if the evaluation systems made a difference in agency achievements. No one can show whether the agency performed better, and whether programs and projects improved. Under current federal policies and direction, it is unlikely that the planning agencies will formulate such objectives.

2. EVALUATION SYSTEMS HAVE AN EVEN CHANCE OF PROVIDING INFORMATION THAT WILL BE USED BY EITHER PLANNING AGENCIES OR CRIMINAL JUSTICE AGENCIES. (Four of eight sites produced such information.)

Half of the eight completed MEP sites demonstrated that they produced information that was considered to be useful. Useful means: (1) the user claimed it was useful; (2) the user cited instances of using the information; and (3) there was evidence that the use cited was important to the performance of the user's function. This means that half the MEP sites were able to fund people who could use evaluation in carrying out the function for which they were responsible.

The limitation of the "use" success criteria for evaluation was raised in Lesson 1 above. One can document if the information was used, but not whether the use was correct or made a difference to the effectiveness of the function, or whether the function is important to the performance of an agency, a program, or a project. Thus, "use" is a necessary but not sufficient condition in judging whether an evaluation system contributes to improved agency performance.

The functions served by the MEP ranged from technical assistance services for subgrantees to county board decision processes. We grouped the functions served into two markets for evaluation information: Project Oversight and Issue Identification. The Project Oversight market consists of SPA, RPU and subgrantees using the evaluation system to design, monitor, manage, and provide technical assistance to LEAA funded projects. This market usually relies on data routinely collected by subgrantees. The Issue Identification market uses the evaluation system to raise, explicate or resolve some issues about policies or programs. For the Issue Identification market, ad hoc, one-time studies are designed to focus on issues considered relevant by some set of actors in the criminal justice system.

As discussed in the next three lessons, the likelihood of evaluation information being used depends on the existence of one of these markets and how it is served. One market appears to be quite common, the other relatively uncommon.

3. EVALUATION SYSTEMS ARE NOT LIKELY TO BE USEFUL TO THE PLANNING AGENCY STAFF UNLESS THE AGENCY IS BEING DIRECTED TOWARD SOME MISSION-RELATED OBJECTIVE. (One of three were useful.)

Three of the MEP sites attempted to provide information for use by the staff of a State Planning Agency. One site demonstrated that the Planning Agency needed and used evaluation information; two demonstrated that the agency did not need or use evaluation information. The successful MEP site was one where SPA management had created a Project Oversight market which used evaluation information to influence criminal justice system actions on the projects funded. The agency activities were directed toward a mission of affecting the operation of the criminal justice system. Although this mission-related objective was not well defined and measurable, the agency had in place staff functions which were expected to and did use evaluation information for the purpose of achieving the objective.

The other two sites attempted to use evaluation information to support their planning, funding and subgrant compliance functions. These functions, as practiced, turned out not to need or use the evaluation information. The functions were not plausibly related to LEAA missions of affecting crime and criminal justice system performance. Instead they had process type objectives: disseminate funds and comply with federal guidelines.

4. EVALUATION SYSTEMS HAVE AN EXCELLENT CHANCE OF PROVIDING USEFUL INFORMATION TO STATE AND LOCAL CRIMINAL JUSTICE AGENCIES IF THEY CHOSE TO SERVE THIS AUDIENCE. (Four attempted and three succeeded.)

Three of the four evaluation systems producing useful information defined the user to be local criminal justice agencies, rather than the planning agency. Apparently there is a significant latent demand for evaluation information and evaluation assistance from this group of users.

These three MEP sites used three different mechanisms to identify and meet the need:

- provide technical assistance upon request to agencies undertaking their own evaluations;
- carry out studies upon request from agencies;
- assign planning agency evaluation teams to user groups which direct the evaluation effort.

At least one other site attempted to serve local criminal justice agencies and failed. The outcome indicates that the demand is not created by simply providing information.

5. AN EVALUATION SYSTEM CAN INCREASE ITS CHANCES OF PRODUCING USEFUL INFORMATION BY: SERVING USERS AND INFORMATION NEEDS THAT ALREADY EXIST, AND ALLOWING THE USERS TO OWN THE STUDY. (Four of eight did so.)

Four of the eight completed sites can be singled out as successful in producing useful information. When these sites are contrasted with other sites, two factors stand out: (1) a demand for evaluation information already existed; and (2) the users felt they owned the studies.

Rather than identify and serve an existing demand, several sites attempted to create a demand for evaluation information by providing it on demonstration or pilot test bases. The evaluators produced information which they thought would be useful to others or would cause others to take action. For example, one site that expected the creation of certain types of evaluation information would cause the SPA to create new planning functions. The evaluators were guessing and usually guessed wrong. In these cases, some users expressed satisfaction with what they received, but few actual uses could be identified.

Having a demand for information is not equated to having an administrative requirement to produce evaluation information. Administrative requirements did not ensure useful information. An existing demand for evaluation meant that people were available to describe the information they wanted and why they wanted it. Such demand was associated with a management systems already in place, or an issue involved with the delivery of criminal justice services.

The most successful evaluation systems allowed the user to own the study. That is, the users felt the study was answering questions they wanted answered, and in a way they understood and agreed with. Ownership required active user direction of, or participation in, the evaluation. When the users had either taken steps to produce the information on their own, or stayed involved with the design and production process, then the information tended to be used.

6. PLANNING AGENCY EVALUATION SYSTEMS CARRYING OUT INTENSIVE EVALUATIONS TO TEST PROGRAM CONCEPTS WILL NOT BE SUCCESSFUL. (Three sites attempted to, but none succeeded.)

Three MEP sites attempted to produce high quality, disseminative information about the costs, effectiveness and impacts of selected program concepts such as Youth Service Bureaus (YSBs) and Special Police Units (SPUs). These systems were designed to develop "knowledge" about program concept in the form of scientifically acceptable research that would be widely disseminated and used in criminal justice planning and programming. Some studies were completed, but none of the sites succeeded in testing the concept. Either the described information was not produced, or the programs were not implemented sufficiently to provide an opportunity to test the concept.

Two characteristics of SPA and RPU projects disable the "knowledge" building evaluations. First, the variation that occurs among projects within program areas makes it difficult or impossible to characterize activities and objectives by similar measures. Second, there is usually found to be an enormous difference between what projects do and what the planning agency thinks they are doing. Consequently, the objectives and evaluation questions the agency is interested in are frequently inappropriate for the projects in place.

State Planning Agencies apparently do not have the resources and capabilities to be developers and testers of program concepts. The cost of evaluation is only one cost involved in successfully doing intensive evaluations. The cost of well designed and implemented projects has to be considered also. The MEP experience indicates that planning agencies and LEAA have underestimated the cost of developing well defined program concepts and getting them implemented in several locations as designed.

B. WHAT EVALUATION SYSTEMS CAN AND CANNOT BE SET UP

The MEP sites were expected to develop evaluation systems tailored to their needs and capabilities. If the tested system appeared successful, they were expected to institutionalize the system by continuation of funding from agency resources.

In all except one case, an evaluation information production process was established. The systems developed varied considerably and demonstrated that a variety of approaches are feasible and appropriate for planning agencies. Moreover, no single production process or methodology assures the utility of information produced.

Of the seven sites that tested a system, three had some success in institutionalizing it. However, even in these cases, the tested systems were significantly modified, and there is great uncertainty whether the continued system will perform as tested.

The experiences of the MEP sites in attempting to develop and institutionalize evaluation systems are summarized in the following five lessons.

1. IF FUNDS ARE EARMARKED FOR EVALUATION, PLANNING AGENCIES ARE ALMOST CERTAIN TO SET UP AN INTERNAL EVALUATION SYSTEM THAT PRODUCES, OR HELPS PRODUCE, INFORMATION. (Seven of eight did so.)

Seven of the eight completed MEP grantees were able to establish a system that produced some evaluation information. In several cases, the systems and products fell short of what had originally been planned. For example, fewer products were produced than anticipated in most sites, some data were not collected and some that were collected were not analyzed, some analyses were not feasible, and certain planned data bases were not created. Yet a majority of the intended products were produced, and it was demonstrated that grants for evaluation systems will produce evaluation output.

2. DEVELOPMENT OF AN EVALUATION CAPABILITY DOES NOT ASSURE INSTITUTIONALIZATION OF THE CAPABILITY. (Five of eight sites not continued)

Five of the eight completed sites are not continuing the evaluation activities operated during the MEP. Discontinuation is due to a variety of reasons, such as: changes in the operation of the planning agency; lack of funding; and failure to implement or build support for selected aspects of the desired system.

Three of the sites are attempting to continue the evaluation activities tested under the MEP and institutionalize them. However, sufficient changes have been made to make it uncertain as to whether the institutionalization will succeed. Reasons for the uncertainty include: changes in key personnel; changing responsibility for key functions; and changes in the planning agency itself.

3. PLANNING AGENCIES APPEAR TO BE UNSUCCESSFUL IN ESTABLISHING EVALUATION SYSTEMS IN LOCAL CRIMINAL JUSTICE AGENCIES. (Four attempts, three unsuccessful)

Four sites planned to develop an evaluation capability in local or state agencies that could function independent of the planning agency. One grant is still under way. The three completed sites were unable to transfer an independent capability to local criminal justice agencies. When their strategies appeared to be infeasible, these sites changed their objectives and set up evaluation systems internal to the planning agency.

Several sites demonstrated that an in-house RPU evaluation system can provide useful evaluation information and assistance to local criminal justice agencies. However, to date, no planning agency has been able to demonstrate that it can cause these local agencies to set up their own evaluation systems.

4. GIVEN THE OPPORTUNITY, PLANNING AGENCIES WILL SET UP EVALUATION SYSTEMS WHICH DIFFER FROM EACH OTHER IN PRODUCT, COST, AND OPERATIONS.

The evaluation systems demonstrated through the MEP varied considerably in how they operated, what they cost, and what they produced. Some produced many products, others produced none, one, or a few. Some MEP staffs performed all the evaluation tasks by following a formalized approach, whereas others were ad hoc, and others took on only a portion of the tasks involved in performing an evaluation, leaving many tasks to contractors or other agencies.

In general, evaluations utilizing complex research designs did not appear. A few studies used comparison groups or pre/post analyses, but most were descriptive or contained planned versus achievement comparisons.

The grantees did not exhibit consistent relationships among cost and products or effects. As could be expected, the sites with more MEP funds tended to directly handle most of the evaluation functions, whereas the smaller grantees had more instances of jointly performing some tasks with local agencies. However, exceptions existed, and no basis was exhibited for an optimum rule of how much funds to make available for evaluation.

5. THE ONLY CHARACTERISTIC OF AN EVALUATION SYSTEM ASSOCIATED WITH UTILITY WAS THE DEGREE OF INVOLVEMENT OF THE USER IN THE EVALUATION ACTIVITY.

The more a user was committed to or dependent on evaluation results, the more utility the study was likely to have. The four sites judged as producing useful information had significant user involvement invested in the conduct of the study.

The users became involved in establishing the evaluation designs, data collection and analysis, and disseminating reports. This involvement allowed direction to be given that ensured the study would better meet user expectations, and frequently led to the user publicly endorsing and disseminating the results to others.

The user involvement does have its costs. Most of the users are state or local officials having several other responsibilities involved in the management of public activities. Their involvement costs them in time expended and recognition given to the effort. But once the effort is invested, they tend to want to realize the results. Involving the user also costs the evaluator time, an investment which can be substantial.

C. WHAT OPERATIONAL PROBLEMS MAY ARISE IN SETTING UP AND OPERATING EVALUATION SYSTEMS

The MEP sites encountered diverse sets of problems in developing evaluation capablity and using it. In some cases, the problems could not be resolved, and it was necessary to change strategies and expectations. In the end, most sites were able to accomodate the problems and produce evaluations. However, delays, cost overruns, and less than desired products often resulted.

Any evaluation manager or practitioner will encounter some of the problems and should provide time and resources to deal with them. The eight specific problems highlighted here are those that seemed to recur most often, or be most troublesome to the MEP sites. The extent they would reappear in some other site is a function of the management, organizational, and environmental characteristics of that planning agency.

There are no general solutions to these problems. When they occurred, evaluation managers developed solutions appropriate to their own situation. These, then, are problems managers can expect to occur, and with which they should be prepared to deal. Many of the problems can be avoided by giving them attention when planning the evaluation system.

1. STAFF TURNOVER IN THE EVALUATION UNIT WILL CAUSE DELAYS AND DISRUPTIONS.

All the MEP sites experienced some staff turnover, and half experienced significant problems. Without staff, the work does not proceed. But even if replacements can be found, problems of inadequate skills or familiarity with the problems at hand necessitate training and time to bring people up to speed. In the MEP these conditions led to delays and abandonment of some tasks. The problems were sufficiently severe that one site was content to operate after the MEP with a smaller staff, instead of attempting to fill positions vacated through attrition.

2. CIVIL SERVICE RULES MAY CAUSE DELAYS.

Several sites had difficulty initiating work due to their inability to obtain staff. Positions had to be authorized and filled through a Civil Service system that moved slowly and did not generate adequate candidates. The effect was to delay work or cause it to be stretched out over longer than planned periods.

3. PROJECTS MAY BE RELUCTANT TO REPORT DATA.

Several of the tested evaluation systems expected the projects to provide data for the analysis by the MEP team. This approach worked well for many projects, but did not work at all for others. Project refusal to report data seems to be part a political problem, and part a resource problem.

Several evaluation offices believed they did not have the authority or management support to require subgrantees to report data. Frequently subgrantees would not report data when they were not required by subgrant conditions. One site had to promise some projects that data reported would not be used in refunding decisions.

The collection and reporting of data is a significant cost to whomever performs it. By transferring the responsibility to the projects being evaluated, the cost was transferred also. Since the MEP sites did not provide additional resources to the projects, the cost was an extra burden often unanticipated by the project. Consequently many projects could not or choose not to cooperate. Some projects stated that the cost of the reporting requirements were beginning to outweigh any expected project benefits.

4. PROJECT FILES MAY CONTAIN LESS DATA THAN EXPECTED OR REQUIRED FOR PLANNED EVALUATION.

Several evaluations depended on existing project files for data. Evaluation analysis plans were developed on the assumption that the data would exist and would readily be available. This frequently was not the case. Either data were missing, or extensive errors, discrepancies, and inconsistencies were discovered. The discovery of these conditions required the redesign of analysis plans, and the allocation of resources to generate other sources of information. The results were delays in and often abandonment of promised analyses.

5. COMPUTER PROCESSING OF DATA CAN BE A TIME-CONSUMING, ERROR-PLAGUED UNDERTAKING.

The sites that attempted to standardize data collection instruments anticipated utilizing computers to process and store data. The computer usage was expected to facilitate rapid and accurate processing of data.

In fact, the opposite occurred. Backlogs developed, processing was not quick enough to return data to projects for verification or timely use, and errors were encountered.

Some of the problems can be traced to the lack of adequate planning and management. Data flows quickly overwhelmed existing capacities to handle the situation. When evaluators began to analyze the processed data, they were dissatisfied with the method used to format and store them. The data were either reprocessed in a form usable by the analyst, or portions of the analysis were dropped.

6. CONTRACTING OUT EVALUATIONS HAS ITS OWN SET OF PROBLEMS.

Four of the twelve MEP sites relied significantly on contractors to develop or implement portions of their evaluation systems, and five others used contractors or consultants in lesser roles. This approach to obtaining evaluation capability avoided such problems as those posed by Civil Service, but still encountered problems of its own.

Several sites had difficulty in obtaining an appropriate contractor quickly. The competitive process requires extensive time and staff resources to implement. Once a contractor is selected, staffing problems are still encountered with personnel turnover and hiring difficulties. Two sites terminated their contractors early due to an inability to obtain the desired products without expending significant staff time. Also, the sites believed the contractors were spending too much time in becoming familiar with agency operations and in documenting what was already known.

If an agency decides to use a contractor, the agency should be prepared to invest a sizeable amount of staff time in working with contractors.

Communication and decision mechanisms between the agency and contractor need to be established. Those MEP sites that did this worked effectively with their contractors.

7. THE IDENTIFIED USERS CAN ADD TO THE COST OF EVALUATION BY HAVING UNCLEAR OR SHIFTING INFORMATION REQUIREMENTS.

Earlier it was pointed out that the ownership or involvement of a user in a study increased the chances useful information would be produced. While use would go up, cost also goes up if the user has unclear or shifting information requirements.

In some cases this problem was experienced when resolution of design issues could not be obtained. In other cases, the problem developed when shifts in requirements caused a change in data collection and analysis activities after they had been initiated. A major problem is that, when given an opportunity to ask questions, most people will do so. The result is a large set of questions, all of which cannot be answered. Not having a mechanism to prioritize questions, many evaluators try to answer most of them, and subsequently spread resources too thin to answer any question adequately or in a timely fashion.

8. DIVERSION OF EVALUATION UNIT STAFF TO OTHER PLANNING AGENCY FUNCTION IS A DANGER SIGNAL.

The evaluation activities supported in most sites were new management functions without a proven record of utility to the agency. When more established functions such as processing grants became understaffed, personnel in the newer functions became targets for filling the gap. In this manner, several sites delayed or diluted planned evaluation activities. Once the momentum was lost, it was difficult to regain.

The sites that were most successful in implementing evaluation systems prevented the MEP Evaluation Unit from diverting evaluation staff to carry out other planning agency functions. This was accomplished either through a strong management process which maintained the integrity of the Unit, or by organizational arrangements which made diversion difficult to initiate. One site made a policy decision during the MEP grant period to abandon in-house evaluators because it could not keep the evaluators from being diverted to other tasks by the Board of Directors. It was believed that less diversion could occur if the evaluators were obtained on a contract basis.

D. THE MEP MESSAGE: USER INVOLVEMENT

The MEP demonstrated that it is easy to get evaluation information produced, but much harder to produce useful information. There are no "right" methods or system; there are no easy answers. Decision makers have to want the evaluation results, and be willing to invest time and resources in obtaining information if it is to have utility. Evaluation cannot make a difference by just existing.

The remainder of this document discusses in greater detail the lessons summarized here. Examples of the sites' experiences are provided to demonstrate and support the points made. Chapter II examines the success of the MEPs in demonstrating that evaluation can make a difference. The strategies used to develop evaluation capability, and the types of system developed are described in Chapter III. Finally, Chapter IV presents the major implementation and operation problems encountered during the MEP. The information is structured to provide guidance to others attempting to establish a useful evaluation capability.

II. SUCCESS OF THE TESTED EVALUATION SYSTEMS

Twelve Model Evaluation Program sites were funded. As of January 1978, eight of the twelve had completed their MEP demonstration; four were still in progress. This chapter discusses the demonstrated success of the evaluation systems tested at the eight completed sites. Section A provides an overview of the program outcomes. Section B presents the three types of outcome originally projected for the MEP that are used here as criteria in assessing the success of the program. The rationale for the criteria and the manner in which they are applied are developed there also. Sections C, D and E assess the performance of the program relative to the three criteria, and Section F gives factors that affected success.

A. OVERVIEW OF THE MODEL EVALUATION PROGRAM OUTCOMES

Three criteria were used to measure the success of an evaluation system: did it help agencies achieve their objectives; did it identify and meet a real demand for evaluation information; and was it continued after the MEP grant expired?

No site was able to demonstrate success on the first criteria. None of the 12 sites attempted to develop or measure the agency objectives, nor were any of the sites able to demonstrate that the evaluation systems helped the agencies achieve their objectives.

Of the eight sites, only one failed to produce evaluation information. Examination of how the information was used at the seven remaining sites led to the conclusion that four sites had demonstrated that they identified and met a real demand for information. Two other sites appeared to be providing a free good service for which there was not a clear need. The remaining site was associated with the production of useful information, but it was not demonstrated that the MEP system was the principal reason the information was produced and used.

The MEP sites were able to develop information for two types of user market--one associated with the design and monitoring of LEAA funded projects, and one associated with identifying or resolving issues about program operations for agency management or policy makers. Four MEP sites attempted to meet a third type of market--developing and disseminating knowledge--but failed to produce the required information.

Three of the eight tested systems were continued initially by the planning agency at the end of the MEP, but the long-term fate of these three is uncertain. The discontinuation decisions at the five other MEP sites were based on a variety of judgments, including: failure to implement desired evaluation system; product perceived as not useful; and severe agency budget cuts prevent continuation.

Figure 1 summarizes the results of the Model Evaluation Program grants. Figure 2 summarizes the level of production and investment among the eight completed sites. The following sections of this chapter define the success

FIGURE 1: OUTCOMES OF THE MODEL EVALUATION PROGRAM

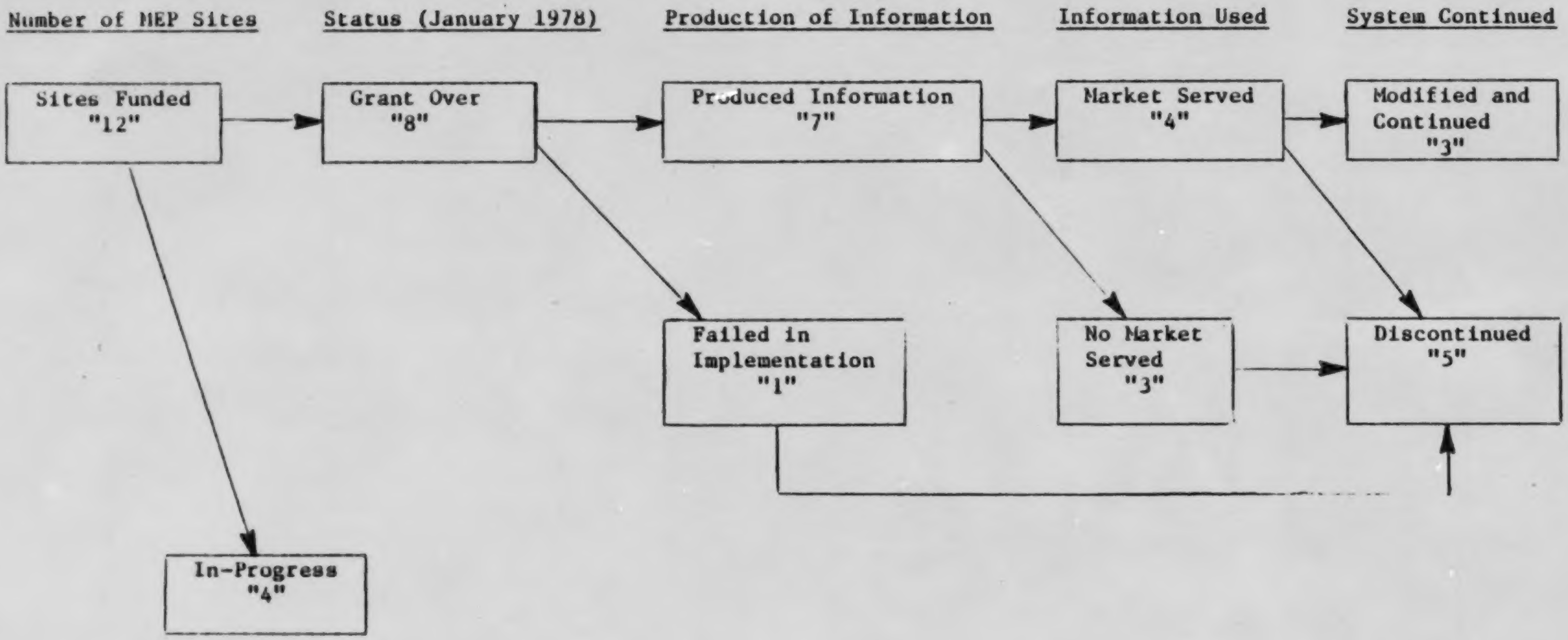

Number of MEP Sites
Status (January 1978)
Production of Information
Information Used
System Continued
Sites Funded "12"
Grant Over "8"
Produced Information "7"
Market Served "4"
Modified and Continued "3"
Failed in Implementation "1"
No Market Served "3"
Discontinued "5"
In-Progress "4"

FIGURE 2: PRODUCTION LEVELS ACHIEVED BY MEP SITES

	MEP Budget* (Rounded to $1,000)	Number of Analyses or Studies Completed	Approximate Cost per Study	Were substantial non-MEP resources used to do the studies?
Central Oklahoma	$108,000	6	$18,000	No
Illinois	$250,000*	17	$14,700	No
Ventura	$ 74,000	18	$ 4,100	Yes
Alameda	$144,000	8	$ 8,000	Yes
Michigan (Reporting System)	$142,000*	4	$35,500	Yes
Michigan (Intensive)	$190,000*	2	$95,000	No
Columbia	$ 56,000*	4	$22,400	No
Jacksonville	$ 85,000	7	$12,140	No
Pennsylvania	$261,000*	0	Not Applicable	No

---

*Two year grants include 50 percent second year match.

criteria, discuss site performance on each, and identify factors associated with success.

B. THREE SUCCESS CRITERIA FOR THE MODEL EVALUATION PROGRAMS: ACHIEVEMENT OF AGENCY OBJECTIVES, USE AND CONTINUATION

The original Model Evaluation Program request for proposal contained three success criteria:[1] (1) evaluation information would be used; (2) evaluation would contribute to the achievement of agency objectives; and (3) the demonstrated evaluation system would be continued if it proved to be useful. At the time of the program announcement, it remained an open question of how to measure and apply the criteria to select grantees or assess success.

The nature of the program required that these criteria be defined as each site developed and implemented its program. At the start of the program, there was uncertainty among all parties involved as to how "use," "achievement of agency objectives," and "continuation" would be interpreted by the MEP agencies.

The evolution of the criteria began with the grantee selection process. Agencies submitting proposals were asked to describe, among other things: their evaluation objectives and goals; their agency objectives; the direct relationship between agency and evaluation objectives; and how the proposed evaluation program could be evaluated. Following grant awards, site visits and interviews were held with the agencies selected to clarify the criteria. As it turned out, none of the sites had developed well-defined, measurable objectives and expectations for their MEP Project or for their own agency. Most, however, had a clear idea of the evaluation system they wished to establish, and how they planned to go about it.

---

[1] Model Evaluation Program Announcement, Washington, D.C., Law Enforcement Assistance Administration, Fall 1974.

The rationale for developing agency success criteria that emerged among the grantees was:

- Each agency planned to produce a specific type of evaluation information for one or more users.
- How that information would be used and the difference it would make could not be predicted by the agency before-hand.
- The agency would produce and disseminate the information and make a decision on continuation based on its experience with the production process and its own estimate of the utility of the product.

Given these conditions, the national evaluation of the Model Evaluation Program adopted a trace method for documenting success. The flow of actual evaluation products was documented and the actions associated with these products identified. Also, the status of the evaluation systems tested as part of the Model Evaluation Program during and following the grant period was traced. This approach enabled the definition of MEP success criteria in a way that reflected both the intent of the program, and what the agencies actually did. Eventually specific measures were developed for the three program success criteria that could be used for all sites.

"Use" of the evaluation information was defined as occurring when three conditions were met:

(1) Intended and actual users claimed the information was useful.

(2) The users cited actual instances of using the information.

(3) There was evidence that the use cited was important to the performance of the user's function.

36

The first two conditions were determined from a survey of recipients of the evaluation information identified by the MEP grantee. Recipients were interviewed to answer such questions as:

- Did they receive evaluation information?
- Did they absorb the information?
- Did it provide information they did not have?
- Did they find it useful?
- How did they use it?
- Could they identify decisions or actions affected by the evaluation system?

Frequently, people interviewed would say the information was useful and used, but had difficulty describing the use. Moreover, it became obvious that differences existed among the cases of use. Some users clearly wanted, needed, and used evaluation information to make a difference in programs and policies. Others were unable to show that the information made a difference, sometimes acknowledging they could get along without it. These latter users appeared to be getting a free good (i.e. evaluation) for which there was no clear need. Therefore, the third condition needed to attribute "use" was established: successful use required evidence that the use cited was important to the performance of the user's function. One of two types of evidence was considered sufficient:

(1) evidence that it made a difference to the outcome of the user's function, such as programmatic or policy changes, enabling a program or policy issue to be raised and debated, or enabling programmatic technical assistance to be delivered; and

(2) evidence that the information was in great demand, such as unusual effort on the part of the user to secure the information, or a formal decision process where the demand for and use of evaluation was highly visible.

The evidence called for by the last of the three conditions would not include such cited uses as: "I found it informative"; "I decided to change my file system as a result"; "I sent the report along to department managers."

The data to apply the criterion of "importance" were drawn from: (a) interviews with users, MEP personnel, and SPA and RPU personnel; and (b) site observations of agency operations, MEP activities, and MEP interaction with users.

The discussion of utility in this report is designed to identify both the uses cited, and to support a conclusion as to whether the MEP served a real market for evaluation, one in which the information made a difference.

"Contribution to the achievement of agency objectives" was undefined at the start of the program, and remained in the end to be undefined. The agencies were examined and agency management was interviewed to determine whether differences in agency performance could be documented and attributed to the Model Evaluation Program. As will be discussed in Section C below, the concept of "agency objectives" has not been operationally defined by State Planning Agencies or by Regional Planning Units, or by the Law Enforcement Assistance Administration. The implications of this outcome are significant. The primary measurable success criterion becomes "use" of information, rather than demonstrable improvement in agency programs or policies.

"Continuation" was measured by tracking the maintenance of MEP resources, and MEP processes and policies after the grant was over. Interviews were held with MEP personnel and agency management to document continuation decisions and reasons for them.

C. MEP CONTRIBUTION TO THE ACHIEVEMENT OF AGENCY OBJECTIVES NOT DEMONSTRABLE

Two of the criteria used to select MEP grantees described the "achievement of agency objectives" as an expectation of the MEP:

1. degree to which the evaluation program goals and objectives are directly related to achieving agency objectives; and

2. likelihood that the evaluation program will help the agency achieve its objectives.

The MEP grantees, then, might have been expected to monitor progress toward overall agency overall objectives and provide information that would enhance agency achievement of those objectives. Most MEP proposals described the agencies' objectives and how the MEP was expected to help achieve them, but they did not provide a basis for management or evaluation of the MEP project or the agency.

Agency objectives were rhetoric-statements without underlying definition or measurement system:

> . . . introducing change and encouraging innovation . . .;
>
> . . . improve the quality and effectiveness of criminal justice programs and projects . . .;
>
> . . . raise the level of consciousness regarding goals and standards . . .;
>
> . . . provide overall coordination . . .;
>
> . . . improve the quality of planning. . . ."

These concepts are not inherently unmeasurable or unattainable; rather, the agencies receiving MEP grants had not defined them.

Grantee objectives for the MEP paralleled those for the agency. They were vague:

> improvement of the comprehensive planning process . . .;
>
> development of priorities and policies for funding . . .;
>
> improve programmatic design . . .;

> increased amount of relevant objective information . . .;
>
> increase awareness (of) the need for evaluation . . .;
>
> provide decision makers with the information they need to implement 'management by objectives (or goals) systems.'

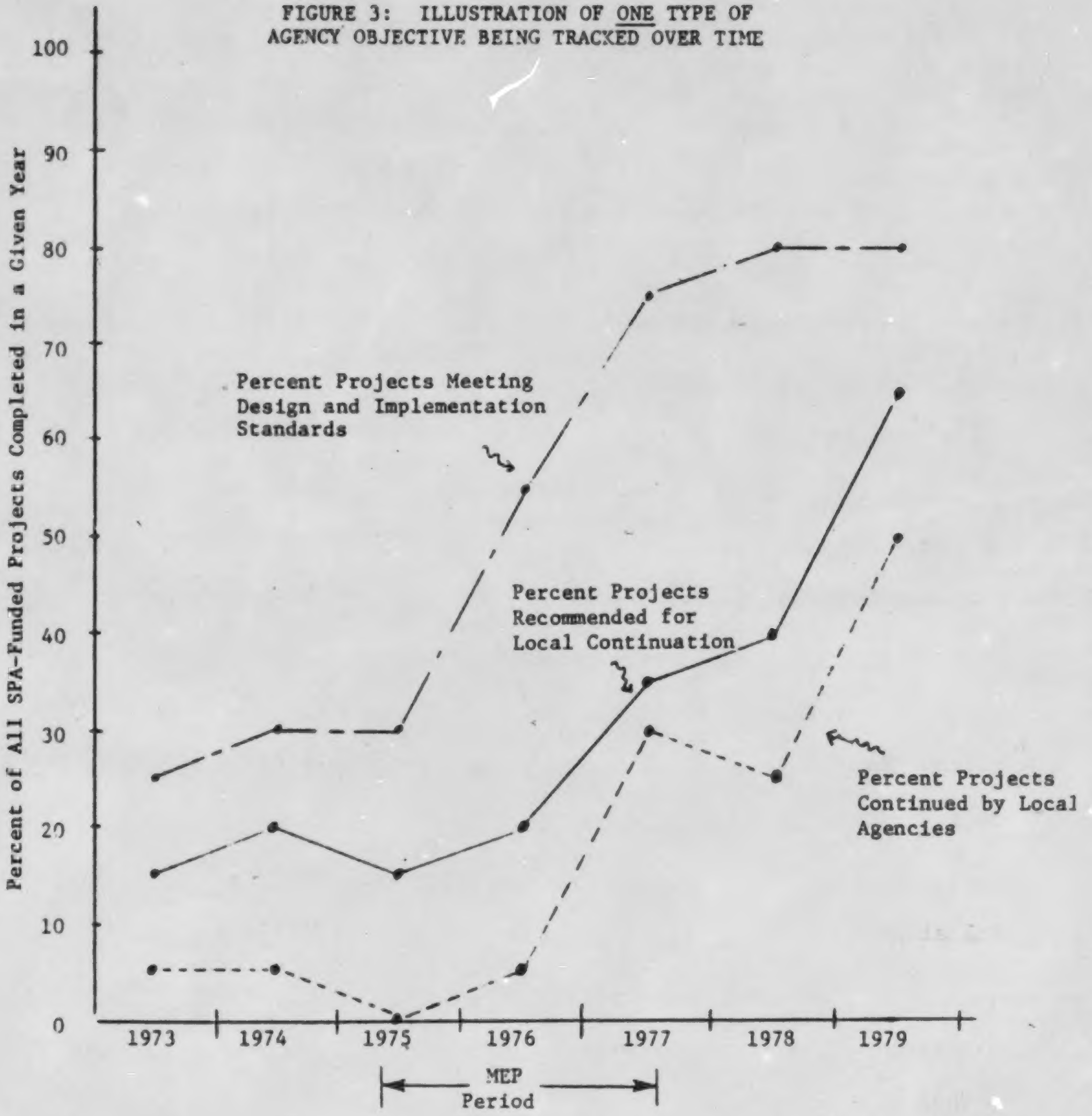
Grantees did not describe how the MEP would help the State Planning Agencies or Regional Planning Units achieve their objectives, or what those objectives were.

The legislative statement authorizing LEAA documents implies that SPAs or RPUs will operationally define such objectives as: improve the criminal justice system, reduce crime, or increase safety. Yet, in neither the agency plans nor agency interviews were descriptions of how the agency would define, achieve, and measure such objectives identified.

This is not to say that the agencies had no objectives toward which they were working. Several appeared to be well administered, with a sense of purpose and direction. The difficulty is that the directives and the objectives were not measurable or verifiable. Performance and progress were subjective determinations made by agency management--"things look like they are getting better," or "this year's plan is better than last year's."

Where agencies appeared to be well administered, it always seems straightforward to translate their rhetorical objectives into verifiable measures. Figure 3 shows how one might define and monitor an objective frequently cited by agencies--"to put in place innovative projects that perform well and are adopted and continued by state and local agencies." Any MEP project could have been designed to improve the agency's performance on the several measures proposed to monitor this objective:

FIGURE 3: ILLUSTRATION OF ONE TYPE OF AGENCY OBJECTIVE BEING TRACKED OVER TIME



— - — Percent projects meeting SPA project design and implementation standards

——— Percent projects that received a "high performance" rating and a recommendation for "local continuation" from a peer review panel of local officials and experts

- - - - - - Percent projects that were continued by local agency at the same or increased operating level

- better project, evaluation and monitoring designs leading to more successful project implementation;
- better evaluations of projects leading to more selective and more persuasive recommendations for local continuation; and
- access to evaluations lead to more continuation decisions.

As Figure 3 implies, if agency objectives were measurable, it might be possible to associate the MEP with "contributing to the achievement of agency objectives." No agency was observed to define and monitor its rhetorical objectives in this way.

If one cannot measure agency performance on its rhetorical objectives, then a dilemma exists as to what objectives, if any, can be measured? Resolution of the dilemma requires one to ignore the rhetoric and examine the activity of an agency to extract a set of measurable objectives from agency actions. Following this line of reasoning, SPAs and RPUs were found to have as an objective the funding and administrative monitoring of subgrantees. That is, they fund subgrantees to carry out projects acceptable to the local criminal justice officials, and in compliance with federal guidelines and regulations. The grants management function may involve technical assistance to subgrantees, as well as monitoring progress and evaluation of results. LEAA requires that some level of evaluation (monitoring or more intensive) be carried out.

When the staff functions of an SPA/RPU have no overall measurable objectives other than selection and funding of subgrants, and compliance to federal guidelines and regulations, one then has an additional dilemma in assessing "the use" of evaluation information. Some means is needed to determine what evaluation information produced by the MEP is supposed to accomplish.

What happened in practice was that the evaluation information from the MEP was disseminated to:

- SPAs/RPUs for use in formal decision processes such as subgrant renewal;
- SPA staff to help carry out their functions such as providing technical assistance and preparing the comprehensive plan;
- subgrantees for their information; and
- criminal justice system officials for their information.

How they were supposed to use it, and for what purpose, were left to the discretion of the user. Success of evaluation on the criteria of achieving agency objectives thus reduces to an assessment of its utility to users.

D. HALF THE MEP SITES WERE ABLE TO DEMONSTRATE THAT A MARKET FOR EVALUATION WAS SERVED

We concluded that four of the eight MEP sites studied demonstrated that a market for evaluation was identified and served. They produced evaluative information, and there is evidence that the use of that information made a difference to the performance of the user's function. Of the four remaining sites, two did not get their information product used; a third left it unclear whether use can be attributed to the MEP; and, in the fourth case, there was no information produced.

The MEP sites attempted to use evaluation to support three different types of decision process (or markets for evaluation):

- Project Oversight: The SPA, RPU, subgrantee process whereby projects are funded, designed, monitored and routinely assessed for refunding decisions.
- Issue Identification: The use of ad hoc studies by decision-making individuals or groups to raise and resolve issues about programs or policies.

- Knowledge Development: The process of testing and disseminating effective program concepts, relying heavily on research standards to produce high quality evaluation information on program cost and performance.

One or more sites were able to demonstrate that useful information could be produced for the first two markets, Project Oversight and Issue Identification. None of the three MEP sites that attempted it were able to produce the evaluation information required by the Knowledge Development market.

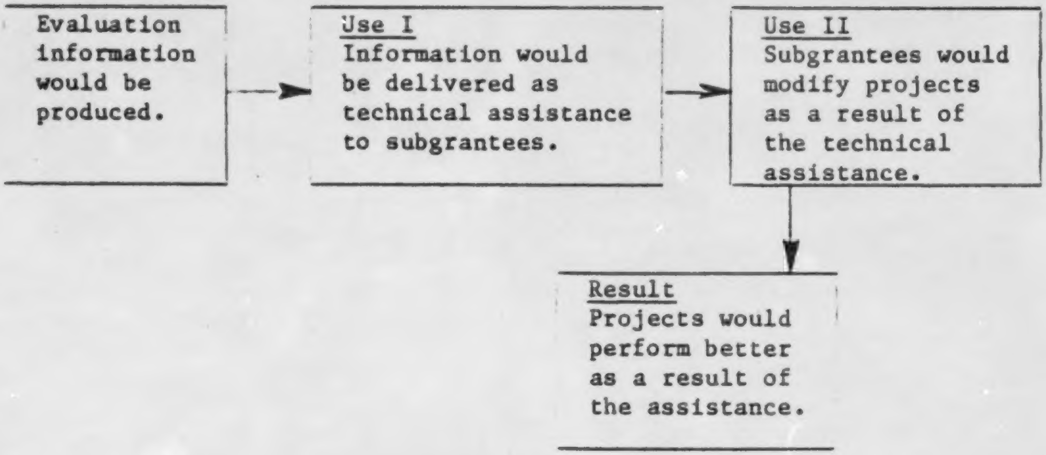
This section describes the limitations of the use criterion, the types of use encountered, and the achievement of the MEP sites on each type of use.

1. SOME LIMITATIONS OF THE USE CRITERION

The evidence of use available to this study has its limitations. This study had to rely heavily on user perceptions of use. The two major problems with this are the reliability of the user observations, and the distance of this indicator from the effect of the information.

When possible, the first problem was dealt with by getting several observers and sometimes several users to describe the consequences of any one study. Our impression is that the observations are reliable--people describe similar things occurring.

The second problem is a more serious limitation. Frequently the "user" interviewed was far removed from the ultimate desired effect of information. For example, one model of use commonly cited assumed the following sequence of events would occur for an SPA technical assistance function:



Evaluation information would be produced.

Use I
Information would be delivered as technical assistance to subgrantees.

Use II
Subgrantees would modify projects as a result of the technical assistance.

Result
Projects would perform better as a result of the assistance.

Clearly, one would like to judge the worth or contribution of the technical assistance function in terms of improvement in project design or performance. As described earlier, agencies were not set up to permit this level of evaluation. Instead, the function of the primary user, in this case deliver technical assistance (Use I), was examined and the question asked: "Did the evaluation system make a difference to the performance of that function?" In other words, could the primary user identify changes or instances of technical assistance attributable to the evaluation? The assessment did not go further to evaluate the technical assistance function itself. Thus, one might say that being successful on the "use" criterion is a necessary, but not sufficient, condition for judging an evaluation system to be beneficial.

2. TYPES OF MARKET FOR EVALUATION

The MEP identified three general types of use for evaluation:

- Project Oversight;
- Issue Identification;
- Knowledge Development.

All represent strategies for improving the criminal justice system. However, each uses evaluation in a different way for a different immediate purpose. Consequently, the uses cited and significance of those uses differ among the three models.

The Project Oversight model uses the evaluation system to help design, monitor, manage and evaluate projects funded by the SPA or RPU. The uses commonly cited under this model are:

- evaluation designs lead to better project designs in subgrant applications;
- monitoring a project uncovered problems which were corrected;
- evaluations affected decisions to refund or continue the project; and
- project performance improved over time as a result of SPA or RPU technical assistance initiated by evaluation information.

The primary users in this model turned out to be one or more of the following: the subgrantee taking action on his or her own initiative; the SPA and RPU staff assigned to monitor and give assistance to subgrantees; the SPA and RPU decision-making bodies deciding on subgrant refunding and subgrant special conditions.

The difference the additional evaluation information would make in the adequacy of project design and in project performance was never described in measurable terms, or recorded by the MEP sites. The success measure for evaluation relied on the primary user's perception of information use, and its influence on project design or performance.

Issue Identification is different from Project Oversight in that: (1) it does not have a role in designing projects and programs prior to evaluation; and (2) it is not an ongoing, routine tracking of some group of projects over time. The Issue Identification model uses the evaluation system to raise, explicate, or resolve some point of an issue about existing or proposed policies and programs. Ad hoc, non-routine studies are designed to meet the situation and focus on issues considered relevant. The uses commonly cited under the model include:

- it showed that a problem existed that had to be solved;
- a policy decision was changed and reconsidered;
- a grant was let to deal with some problem identified;
- an agency was reorganized; or
- a program was redesigned.

The primary users are the decision-making individual or group who acknowledge that an issue exists or must deal with it.

The Issue Identification model attempts to raise and resolve issues. If many of these studies were done and nothing wrong was ever found, then, at some point, one would have to question their utility. (The Project Oversight model, on the other hand, seeks to avoid issues by careful planning and monitoring, and the early correction of potential problems.) The success measure for evaluation in the Issue Identification model relied on primary users' perceptions of the impact of information on their attitudes and on their decisions. The difficulty with interpreting user opinion in this Issue model is that anything can be an issue, and the cost of raising them is not borne by the user. Issues identified ranged from "project files are not adequate," to "should we build a prison." Therefore, evidence was required that the user considered the information to be necessary, and the issues were significant to performance of the users' responsibilities.

The Knowledge Development model uses the evaluation system to produce and disseminate high quality, disseminable information about what programs work best (or worst), under what conditions. Whereas in the other two models the information addresses specific projects or issues, in the Knowledge Development model, general statements about the cost and performance of some program concept are generated and supported by scientifically acceptable research. The uses commonly cited under the model include:

- the identification of optimum program models;
- setting of funding priorities among programs in the Comprehensive Plan;
- dissemination of program and project models to state and local agencies.

The primary user was usually identified to be the SPA or RPU in either a dissemination role or in setting its own funding policies. Success indicators for this model would be the number of "practices" conclusively evaluated, and the number of decisions influenced by that information. In the MEP no one was able to successfully implement this model.

It should be noted that a less rigorous type of knowledge development occurs in the other two models of use. SPA and RPU staff can develop knowledge about what does and doesn't work through their experience with projects. The knowledge will be softer by research standards and less accessible, but it can be equally valid and useful.

Table 1 shows for which markets the sites planned to develop information, and for which markets they actually produced information. Table 2 summarizes conclusions about which sites produced information that was used, i.e., which sites identified and served a real market for evaluation.

3. PROJECT OVERSIGHT

Four sites attempted to serve the Project Oversight market; only one,

TABLE 1: TYPE OF USE FOR WHICH MEP SITES PLANNED TO DEVELOP INFORMATION AND WERE ABLE TO DEVELOP INFORMATION

MEP Site	Project Oversight	Issue Raising and Resolution	Knowledge Development
Central Oklahoma	X	✓	
Illinois	✓		X
Pennsylvania			X
Ventura		✓	
Alameda		✓	
Michigan (Reporting System)	✓		
Michigan (Intensive Evaluation)		✓	X
Jacksonville		✓	
Columbia	✓		

✓ - Developed information
X - Planned to but did not develop such information

TABLE 2: SUCCESS IN PRODUCING USEFUL INFORMATION

For Project Oversight Market

Illinois:	Some information use, but not critical to function of users, and no significant results from use.
Columbia:	Isolated cases of uses that affected the criminal justice system. Doubtful if use can be attributed solely to MEP.
Michigan:	Considered successful, with uses cited.
Central Oklahoma:	Information for this market not produced.

For Issue Identification

Central Oklahoma:	Users claim information useful, but few uses cited, and none critical to function of user.
Ventura:	Successful information used to assist user perform criminal justice functions.
Alameda:	Successful. Uses cited that were significant and depended on the information produced.
Michigan:	Successful. Information used to alter management and program operations.
Jacksonville:	Users satisfied with information and uses cited.

For Knowledge Development Market

No site developed information for this market.

Michigan, demonstrated that a real market for evaluation information was identified and served by the evaluation system.

Three of the four sites were able to develop information to test in the Project Oversight market: Michigan, Illinois and Columbia. Michigan and Illinois tested standardized SPA project reporting systems. For such systems, all projects funded under a "Program Element" in the State Plan would be defined and measured using a common set of measures developed by the SPA. Columbia (RPU) used a different approach--a regional evaluator to help subgrantees design projects and evaluations to meet SPA requirements. A fourth site, Central Oklahoma (RPU), developed a project design methodology for use by SPA subgrantees and local agencies, but did not test it. Here we discuss the use of the information produced at the three sites that did produce information for the Project Oversight market.

a. Michigan (Reporting System)

With its "standard evaluation system," Michigan demonstrated that SPA management can create a demand for evaluation information within the SPA. The standard evaluation system developed process and outcome data on similar projects using common, standard reporting instruments. The system was pilot tested on 20 projects in 4 program areas. It was to produce 20 project analyses and 4 multi-project analyses on the program areas.

At the time this report was written, Michigan was still carrying out the computer analyses and experiencing some difficulty with the com--puter contractor. However, the SPA program staff were familiar with the designs and the data, and had experience in using similar types of information in the past. Based on the work done to date, observation of

Michigan's decision processes and interviews with the primary users, we judged the system to be successful on the use criteria. Our observations are described below.

(1) Users Perceived the Information To Be Useful: At the end of the grant, we interviewed four program specialists and asked them to describe how useful they expected the information to be, and how they expected to use it. They responded that the information was necessary and would be useful. They identified three ways in which the information would be useful. It would help their monitoring function:

> . . . hope to make overall conclusions about value of a program element. . . . Find gross conclusions on what stands out, and go look at that. Decide what should be added to or dropped from projects;

help them justify programs and argue for continuation locally:

> Confirm what we have endorsed and encourage locals to put up dollars;

and lead to changes in project design or operations:

> (In the past) used on individual basis to improve projects . . . and in setting refunding requirements.

(2) Users Cited Instances of Using This Type of Information in the Past: The specialists could identify having used similar information for these purposes in the past. For example, they had changed the design and operation of programs.

> For jail rehabilitation programs, their budget is sent in with staff salaries and what they will do. But in reality it takes six to seven months to get going. During the time, they mainly staff up and find a home. As a result of the review of data, we now first fund them

> for one to six months to perform the administrative work--hiring, etc. Then we fund for a period of providing services.
>
> One methodology for a type of program appeared superior at one-fourth the cost. As a result, we revised the (alternative education program). . . . found that lower pupil/instructor ratio and taking pupils to alternative educations . . . proved to be effective.

Each of the program specialists, when asked, held that the standard evaluations were necessary, and that they should continue to design such evaluations for other projects and programs.

(3) There Is Evidence That the Use Cited Is Important to the Performance of the Market Identified: Four observations indicate that there exists a real demand for the evaluation system: SPA management's investment of its own resources into evaluation, SPA management's interest and support of the system, the existence of a formal decision-making process which uses evaluative information, and the support of the data collection by RPUs and projects.

The Michigan SPA demonstrated that the information was wanted by the commitment of management time and its program staff time to evaluation activities. Management invested a significant amount of program staff time (that is, non-evaluation staff) in training to increase staff evaluation capability. The program specialists are also required to allocate their time on an on-going basis to the design and implementation of standard evaluations for their projects. The level of involvement of non-evaluation staff personnel is exceptionally high in comparison to other SPAs and RPUs.

During the course of the MEP, we interviewed six members of the SPA top management, and all were aware of the MEP, supportive, and participating in it. Moreover, they all were expecting improved performance from the program staff in terms of well-designed projects and better staff assessments and recommendations. Management demonstrated its support for the system by demanding participation in the training program when attendance began to slip.

The need for evaluation arises by the unique decision-making process that exists in Michigan, and the role the program specialists play in it. The Michigan SPA management is unique in that its governing Commission has delegated a great deal of policy-making and funding authority to the SPA management. Thus, the Director and the line-supervisors constitute an active and strong decision-making body. For planning and grant review sessions, they have required that the program specialists bring forward evaluative information on projects and programs. In addition, the program specialists are responsible for giving technical assistance to potential applicants and subgrantees on project design and operations. Interviews with six of the sixteen program specialists confirmed the existence of these roles. The program specialists used the standard system to help define projects, to help monitor projects, and to accumulate expertise on programs.

Finally, the response of other non-SPA actors lends credibility to the system. The Michigan MEP invested time in meeting with regional staff and project staff to review and discuss the standard instruments and reporting requirements. Moreover, site visits were made to each project to verify the applicability of the instrument and the project's ability to collect the data. The recorded comments of regional and project staff indicate that the evaluation system was viewed as necessary, appropriate and useful. Subsequent data reporting was generally good.

(4) The Information Used Is Attributable to the MEP Evaluation System: The MEP standard evaluation is, in effect, a further development and improvement of an evaluation and management system that already existed at the SPA. Our interviews and reviews of site material indicate that the MEP would make a significant difference in the quantity and quality of project information flowing between subgrantee and program specialists.

Prior to the MEP, projects were required to report summary data on standard factors, but the requirement was not enforced or monitored. Reporting was spotty, and reports varied in quality and comprehensibility. The reports from similar projects were not comparable. With the standard instruments, enforcement, central analysis and cross-project analysis, the specialists were expecting more detail, more questions answered and clearer answers.

b. Illinois

The Illinois project reporting system did not demonstrate that it produced information that was used. This system was similar in form, but not in operation, to the Michigan reporting system. Standard reporting instruments were used by subgrantees to routinely report data which the SPA would process and analyze. The Illinois system developed evaluation packages estimated to provide data collection instruments for about 60 percent of the funded projects. The system produced 16 project analyses and 1 multi-project analysis during the MEP to test the use of the product.

The primary users were the SPA grants management and technical assistance staff, and the SPA evaluation staff which participated in some reviews of applications for continuation funding.

(1) Usefulness of the Information Was Questionable: Five program personnel (grants managers and technical assistance staff), seven evaluators and two SPA management personnel were interviewed. The general view of the evaluation information was: (1) that it was not used to monitor projects; and (2) that it may have been useful for grant review committees, but there were other, equally significant sources of information.

Four of the five program staff described the information as not useful:

- One was not aware of the information he was receiving for his project areas. Over the past year, he had not used evaluation office products, or had much contact with the evaluation office staff.
- Two stated that they receive the evaluation data submitted by projects, but do not read or file them. They asserted that they did not use the data for managerial purposes. When refunding decisions are to be made, they will read the results of an evaluation if they are available.
- One claimed that he does not use the evaluation data for monitoring purposes. Monthly meetings with projects and telephone calls provide a better mechanism for feedback and discussion.

A fifth person claimed that he did use the evaluation reports as a basis for developing special conditions on some grants. Similarly, four of the seven evaluators identified the grant review process as where the information was useful.

(2) Citations of Use Were Few: Special conditions on grants and project guideline development were cited as the primary uses:

> developing special conditions . . . such as ensuring that the intended target population was being served;
>
> as a result of some informal analyses with the data . . . original guidelines on what constituted a deferred prosecution project were modified.

These were the only specific uses directly attributed to the information. While further probing might have uncovered others, most people interviewed implied that the information produced was not essential.

(3) The Evidence Was that the Use Cited Was Not Necessary to the Performance of the Market Identified: Several observations led one to question whether the information produced made a difference or was needed: There was no sign of a great demand among program personnel; the claim that the evaluators and the evaluations contributed to the grant review process was not supported by management; and the casual implementation of the system makes it suspect.

Interviews with program staff produced no evidence that they required or wanted the information. Most seemed to have their own sources of information on which they would rather rely. They did not invest significant amounts of their own time in developing or using the information.

Neither of the two management level persons interviewd claimed that the system made a difference after a year of operation. When asked if MEP had made a difference and its products used, one SPA manager stated that:

> He did not think there had been use, except to respond to a specific LEAA request for data. He found it hard to differentiate whether there had been any change over the year in the amount of data that existed. He stated that there had never really been a problem with the lack of a data base, even though data

> deficiencies did exist on occasion. However, there was not such a marked change in data quality that it was noticeable to him.

The other manager interviewed could not identify any payoff, other than more detailed project descriptions in the 1977 comprehensive plan.

Management also claimed that the inclusion of the evaluators in the grant review process, the primary use cited by evaluators, just did not work:

> The staff was already over-committed and did not have time to allocate to the activity. Also, the project designs were of such a poor nature that the evaluation review was not really feasible except on a superficial nature.

However, one manager judged the effort of producing the evaluation information worthwhile, since it did lead to "some special conditions aimed at tightening up the projects." The attribution of this effect to the MEP was more uncertain to the other manager interviewed. When recommendations were made to the Executive Committee concerning a program or a project, he would not know the source of the information used to support the recommendations.

Finally, the evaluation system was implemented in such a way that it is difficult to take its products seriously. Many projects were told that the data would not be used in making refunding decisions. For the data that were submitted, there was no quality control over the data, or attempts to correct errors. Data or analyses were never fed back to the projects for review or validation. There were no analysis plans or routines developed for much of the reported data. In short, the quality, validity, and utility of the product are suspect.

(4) Information Use Is Attributable to the Tested Evaluation System: The information produced and the use that did occur are attributable to the evaluation system tested as part of the MEP. The reports that were produced were obtained from data collected with the evaluation packages developed for the system, and through analyses conducted as part of the system.

The system itself is only partially attributable to the MEP. The Illinois SPA had begun development and operation of the evaluation system using standardized data systems prior to receipt of the MEP. In fact, 17 of the 36 evaluation packages eventually developed were developed or being developed prior to the MEP grant award, and reports produced developed during the MEP utilized the pre-MEP data collection instruments. The MEP grant served to accelerate the development process and provide some activities not previously planned.

The limited level of operation of the system achieved is, therefore, not solely due to the MEP. However, the number of standardized evaluation instruments, the amount of data collected, and the number of reports produced would not have occurred without the augmentation with MEP funds. Furthermore, operation of the system to the point of breakdown would not have been realized. Thus, much of the information produced and lessons learned are due to the MEP; the overall system cannot be attributed to the program.

c. Columbia

The Columbia MEP was associated with information that was used by both the project directors and the SPA. However, it is unclear whether the production of the information can be attributed to the MEP. Our observation of the activity of the MEP and interviews with participants raise doubts about the MEP contribution.

Columbia established an evaluation assistance capability rather than a set of designs and procedures as did Michigan and Illinois. An RPU position was established with the MEP grant to help local subgrantees meet SPA evaluation requirements and, at the same time, to develop information that they, the subgrantees, could use. This facilitator was involved with eight projects over a two year period. At the end of the grant, six people responsible for six different projects and SPA evaluation personnel were interviewed.

(1) Subgrantees Perceived the Information To Be Useful: Four of the six subgrantees interviewed claimed they had used the information in decision making; five of the six said they plan to use the information.

(2) Subgrantees Primarily Cited Examples of Intended Use: One subgrantee claimed to have modified his project based on information developed in the evaluation:

- Crime trend data developed by the study were used to shift police manpower and equipment (participating in the project).

Three of the subgrantees cited how they planned to use the information in program planning and design:

> Information changed our way of thinking. Now more prevention oriented. . . . Plan to use it to decide where to set up preventive patrol.
>
> Plan to use geographic data of study to plan new programs. Found that largest single group (of clients) comes from public housing; therefore, will set up satellite agency in those areas.

> Using the results of the utilization of men and equipment in this study to find a better procedure to allocate men and equipment.

A fifth planned more detailed evaluations as a result of the experience.

> We have redesigned the evaluation after looking at the results. Now changed to collect more detailed data. Will need 12 months before can use data.

(3) There Is Evidence that the Uses Cited Are Important to the Market Identified: The market for evaluation in Columbia is created in part by SPA requirements that evaluations be done, and in part by the interest of the subgrantees to use information.

There is strong evidence that the SPA actually reviews and uses the information it requires projects to produce. Early in the grant the MEP evaluator prepared evaluation progress reports for subgrantees that directly led to the SPA cancelling two projects. This action caused the local agency involved to take the evaluation requirements and the content of reports more seriously.

As the MEP work evolved, the project personnel were investing their own time and effort in evaluation activities. Four of the six claimed to have developed positive attitudes towards evaluation as a result of the MEP experience. For example:

> Now I feel good about evaluations and can see what they do and what impact they have.
>
> See accomplishments of an updated record system.
>
> Now we can tell where we were, what we did, and where we are going.

Two respondents indicated no change in their opinions of evaluation--one had been positive, the other negative, prior to the MEP.

Much of the MEP effort went toward working with one agency (Police Department) on two of its four subgrants. The SPA, the MEP evaluator, and the Police Department acknowledge a positive change in the agency's attitude towards, and investment in, evaluation. However, this was the agency that saw two grant cancellations associated with evaluations.

(4) While Information Was Generated and Used, It Is Not Clear that the MEP Was Required for this To Have Occurred: The production and use of information cited could be attributed to the SPA evaluation system. There is no clear cut evidence that the RPU evaluator did or did not make a difference to that system.

Four of the six subgrantees believed that the evaluation activities carried out by their projects could have been done without the regional evaluator. The other two view the regional evaluator undertaking research-type activities they would not have had the skills or resources to do.

The role of the MEP evaluator most frequently identified by local personnel was one of answering questions and giving advice concerning the evaluation requirement of the subgrant. One project claimed there was no involvement worth mentioning, whereas three projects identified involvement in setting up data collection and analysis procedures, and some participation in the implementation of the procedures. Four of the six persons interviewed identified having someone to answer evaluation questions as the main advantage of having a regional evaluator.

The SPA evaluation director claimed that, from the SPA perspective, the regional evaluator may not have added to the information available, but that the MEP appeared to have helped institutionalize evaluation activity in the Columbia Police Department. As indicated earlier, following receipt of quarterly reports, the SPA cancelled two of the grants and required a redesign of the third. The reports had been prepared by the regional evaluator. The SPA staff claimed that the situation would eventually have come to the attention of their monitor and resulted in similar action without the MEP evaluator's activities. That is, the timing may have been accelerated, but the actions would eventually have occurred. The MEP evaluator did not share that opinion.

The RPU, in the final progress report to the National Institute for the MEP project, cites the changes in the Columbia Police Department as evidence of achievement of the objective to improve local attitudes toward evaluation. An SPA evaluator concurs in this view by noting that without the presence of a regional evaluator working with the Department on SPA evaluations, its acceptance and use of evaluation may not have occurred.

In summary, the regional evaluator appears to have had an influence on two subgrantees. However, overall, the uses cited are tentative, and the role of the evaluator is not consistently a significant one across projects. Consequently, the experience at the Columbia site is judged to be inconclusive.

4. ISSUE IDENTIFICATION

Four sites demonstrated that they were able to produce information used in the Issue market. In all, five MEP sites attempted to serve the market. Michigan, Ventura and Alameda were successful at producing useful information. Jacksonville had some evidence of success, but it was not as strong as the other three sites. The Central Oklahoma effort was judged not to be successful in serving this use.

All of the five sites produced a number of separate studies which raised issues about some criminal justice program or project. Each site differed in approach and type of study:

- Michigan: Carried out under contract two intensive evaluations covering two program areas and involving forty-three projects.
- Ventura: RPU evaluator gave technical assistance to 18 local agency evaluation efforts, most of which would not have been done or have been as useful without this assistance.
- Alameda: Eight studies in two program areas were carried out by RPU staff, with control of the studies delegated to local agency officials designated as users.
- Jacksonville: Seven studies were carried out by RPU staff, some upon local agency activities, some on RPU subgrants.
- ACOG: RPU staff produced six individual project evaluations using a common evaluation design approach.

The following is a discussion of how this information was described as being used.

a. Michigan

The Michigan SPA produced two intensive evaluations on two program elements covering forty-three projects. The users of the information were the SPA top management, and its program staff responsible for the program element. Five members of a nine member management group and two program specialists were interviewed.

(1) Users Perceived the Evaluation as Being Useful: All seven users were familiar with one or both of the studies. When asked, they were able to cite the study findings.

All but one person responded that: (1) they believed the study findings; and (2) they felt that they had received new information. The new information was described as:

- Special Police Unit Study (SPU)

  SPUs are not as effective as program specialists have implied; program specialists get to be proponents.

  SPUs did not do what they said they would do.

  Information on project implementation strategy and problem gave us statistical analysis on the way SPUs function, which we did not have before the study.

- Youth Service Bureau Study (YSB)

  Surprised that impact on crime not there.

  YSB not doing what we thought.

  YSB not a unitary concept; are all things to all people; no commonality of business.

Respondents claimed that this information changed or reinforced their attitudes toward the program.

(2) Users Could Cite a Number of Organizational and Policy Changes under Way or under Discussion: Several respondents said that the two studies led to a major change in SPA policies regarding how program staff spend their time. The SPA has put new emphasis on program specialists spending more time in the field, and giving technical assistance to subgrantees during the proposal writing and project design phase. The two studies had concluded that the projects were not operating as assumed by the SPA, or as planned in grant proposals. Interviews with the specialists confirmed that these policy changes had been issued.

Each user responded to the use question differently. In the aggregate, Michigan has taken or plans to take the following types of action based on the study:

- Change the SPA planning and funding process or criteria:

  Change or tighten proposal requirements.

  . . . change requirements before new grantees can submit proposals.

  . . . in administrative review (we are) using the studies to ask applicants if they are doing what (they planned).

- Change how the program specialists spend their time.

  We now have program specialists in the field to participate (with subgrantee) in (project) design.

  . . . encourage YSB staff and program specialists to do a better job in tracking and monitoring.

  . . . spending more time in the field . . . there is an office policy to give more technical assistance . . . give technical assistance before funding and help clean up the details.

- Provide technical assistance to subgrantees.

  Used results in technical assistance discussions with YSB directors.

  We are funding a project to give technical assistance to all state YSB projects, and will use this study and lessons from this study.

- Disseminate the information.

  Wrote a new section for the Commission on youth diversion . . .

. . . sent it to state YSB organizations.

. . . discuss results with projects

- Discuss the implications of the report in the management group.

Used it in internal discussions at the SPA.

Used it in discussions, but not in decision making.

(3) There Is Evidence that the Use Cited Is Important to the Performance of the Market: Management's reaction to the findings and investment of non-evaluation staff time in the effort is evidence that the study is making a difference.

The changes made by Michigan in organizational policies indicate that the results of the study are being taken seriously. The study results were overwhelmingly negative:

- There was no evidence of effects on crime and on the criminal justice system.
- Projects were never implemented as planned or expected.

The contractor concluded "that the apparent lack of project effects is due to implementation failure, rather than conceptual failure." These findings caused Michigan SPA management to re-examine its administrative policies and change them. Michigan plans to go ahead with future intensive evaluations, but with tighter project designs and SPA monitoring of implementation.

The Michigan SPA also devoted a relatively significant portion of management and program staff time to guiding and reviewing the implementation of these studies. While the involvement was not as big as with the standard evaluations, there was participation by non-evaluation personnel.

(4) The Information Used Is Attributable to the MEP Evaluation System: The two intensive evaluations were funded directly out of the MEP grant. The contract was a pilot test of both the utility of intensive evaluations, and the phased contracts used to carry out the studies. The production of the information and its utilization are directly attributable to the MEP.

b. Ventura

The Ventura RPU demonstrated that there existed a demand for evaluation information among criminal justice agencies, and a market for technical assistance in evaluation.

By providing technical assistance in evaluation, the Ventura MEP was involved and partially responsible for the production of 18 studies. The users were officials in local agencies.

(1) Users Perceived the Evaluation Information To Be Useful: Eleven local officials were interviewed concerning the usefulness of evaluation results. Several were users of more than one study, so the user sample is 14. Of the 14, 13 thought the information produced was useful, and only 1 remarked: "We could have gotten along without it."

(2) Users Claimed the Studies Were Informative, and about Half Could Cite Program Changes that Resulted: The following selected quotes and descriptions illustrate the comments of the users.

- Informative (six of thirteen)

  (Drunk Driving Study) It put to
  rest the argument that we ought

to do something different than the way we're doing it now. The study validated our general belief that there didn't need to be a change in our policy with regard to the drunk driver--District Attorney.

(Jail Study) It helped in overall planning--Police official.

(Sentencing Pattern Study) Contributed to more knowledge based on hard data of what is going on, and what has been going on. The problem is that the final report has not been published, and relatively few people have read it --Judge.

Three respondents commenting on the same study indicated that it is affecting opinions, and that they could have used it; however, at the time of the interview, the State Attorney General had ruled that that type of program could no longer be operated. So the study has been shelved until such time as the Attorney General's ruling is changed.

- Changes in Operation (seven of thirteen)

(Witness Utilization Study) We adjusted our calendars so as not to hold officers unnecessarily. Also introduced a one day delay in jury trials to reduce officer overtime --Judge.

(Organizational Study) A few times had thought that a couple of supervisors (officers) were a little arbitrary with the troops, and that some team training or supervisory training might be necessary. But after reading the report, decided that didn't need to do this--Police Chief.

> (Victimization Study) Some of the things learned caused us to institute a burglary task force at Christmas time--Police Chief.

(3) There Is Evidence that the Uses Cited Are Important to the Performance of the Market Identified: Three observations indicate that there existed a real market for the evaluation information: initiative of the local agencies in undertaking these evaluations; the investment of local resources in doing the studies; and the nature of the uses cited.

Most of the evaluations which the RPU participated in were local agency efforts undertaken at the initiative of the agency. The agency usually had invested its own resources and carried out many of the necessary steps on its own. The MEP was in a supportive role, providing resources and/or technical assistance and/or services.

The Ventura RPU added to the quantity and quality of evaluations that the agencies wanted to do. Having the evaluations underway without the MEP indicates that they wanted and needed the information. Moreover, the user interviews indicate that when local officials received the information, they used it to change policies and operations. All users interviewed were familiar with the results of the study, and most could cite planned or actual uses.

(4) The Information Used Is Attributable to the MEP: Since Ventura was acting in a technical assistance role, one might question whether the MEP really was necessary. Based on the user sureys and observation of MEP activities, we concluded that the Ventura MEP was responsible for much of the information produced.

Eleven of the fourteen users said that the MEP enabled a study to be done that would not have been done without the MEP. Ten claimed that the MEP involvement improved the quality of the study. In many cases the MEP evaluator provided resources or assistance which the projects probably could not have gotten easily on their own initiative--e.g., computer time, programming assistance, and analysis capability.

c. Alameda (Detention Studies)

The Alameda MEP experiences demonstrate that a market for evaluation information exists among local officials that can be served through a Regional Planning Unit. The Alameda RPU carried out two sets of studies during the MEP which involved two different mechanisms for serving the market. In one case, the market could be served through existing local organizations: evaluation information was provided to a local decision-making body on controversial issues surrounding the construction of new detention facilities and the operation of pre-trial detention programs. In the other case, a mechanism for serving the market was established: the RPU brought together local officials concerned about an area of criminal justice; provided them evaluation information that identified issues; and they lobbied for action on the issues. Experiences with the detention studies are presented here. The following subsection, (d), discusses the juvenile studies.

The studies of pre-trial detention services and facilities were conducted for the Alameda County Board of Supervisors. The Board and County Administration were to make decisions on the future of the services and facilities, and issues existed on which there were recognized information gaps. The RPU offered to fill the information gaps with studies of the issues. The Board publicly approved the RPU performing the studies and received two completed reports at a public hearing.

One report addressed the jail facility issues; the other the operation of the pre-trial service activities. Subsequently, staff and the public reviewed the reports, and actions were taken on the issues that had been studied.

(1) Local Officials Perceive the Detention Studies As Being Useful: Ten local officials were interviewed concerning the utility of the detention studies, and documents from a dozen other sources were reviewed. Information was obtained from County Board members, County Administrators, local agency heads and citizen groups. All evidence indicates that the evaluation studies were used and considered useful.

All persons interviewed were definitely aware of the studies, had read them, and had used them in developing positions or action plans. A member of the County Board claimed that the study convinced him that plans for selected jail facilities could not be delayed due to the performance of pre-trial services. Another board member cited the provision of additional information and options for planning jail facilities as the primary uses of the information.

Four of the ten persons interviewed cited the studies as a catalyst to decisions on reorganization of the pre-trial services organizations. They claimed it was known that changes would occur soon, but the study helped focus attention on issues and provide an information base for acting. The reverse finding was true on the detention facility studies. Four officials stated they believed that the debates that preceeded a decision had been extended by the introduction of information from the study. "It opened up for debate options that had been resolved long ago, and thus delayed inevitable decisions. In this regard it was a disservice.'

(2) Users Cited Instances of Using the Detention Studies: The users of the study on the pre-trial services cited the study as the catalyst that led to the reorganization of the Division of Pre-trial Services. Various organizational schemes were debated, judges and other officials took public positions, and a new scheme was eventually adopted by the County. While the change was attributed to the study, the modifications did not reflect the study's recommendations. The new organizational arrangement included the transfer of the County Office for Coordination of Pre-trial Services from the Probation Department to direct support of the courts.

In the case of the Detention report, the County Board was committed to make a decision on new jail facilities for pre-trial detention. The questions to be decided involved the size and configuration of the facilities. After publication and review of the report, a decision was made that differed both from the position previously recommended by the County Administrator, and the options given in the RPU's report.

(3) There Is Evidence that the Use Cited Is Important to the Users' Function: The demand exhibited for the evaluation information provide evidence of the importance of the information to the County government and the resolution of the issue. The studies were initially requested by the County Board, and the Board prior to receipt of the MEP had authorized funds for conduct of a smaller study in the same area. County funds were eventually utilized to complete the studies.

All persons interviewed concurred in the point that the studies had an effect in changing the organization of pre-trial services and the outcome of the detention facility decisions. The changes did not follow the MEP recommended actions, but were aimed at resolving the issues studied and raised.

(4) Information Used Is Attributable to the MEP: Local officials had acknowledged an information gap prior to the MEP and had initiated actions to fill it. However, the information production process was assumed and expanded by the MEP. In particular, the expansion of the study effort to the organization of the pre-trial services is directly due to the MEP.

d. Alameda (Juvenile Studies)

The studies of the juvenile justice system were conducted under the guidance of a Juvenile Justice Committee of citizens and local officials established by the RPU. The Committee was charged with studying juvenile justice practices in Alameda, and using the study results to define model juvenile justice practices and get them adopted. A series of five studies were completed and provided to the Committee for use in defining issues and taking actions on them.

For this set of studies, the time invested by local officials in evaluation supervision, review of findings, preparation of recommendations, and lobbying, indicates that there was a serious need and desire for evaluation. Also, local officials were able to identify changes in policies, funding decisions, and operations resulting from the studies.

(1) Local Officials Perceived the Juvenile Studies As Being Useful: Six members of the Juvenile Justice Committee and six local officials working in juvenile justice were interviewed concerning the utility of the studies. All interviewed were familiar with the studies, and could describe how they used the studies. All claimed that the studies were useful.

The six interviewed members of the Juvenile Justice Committee used the studies to develop recommendations for changes in the local juvenile system, and to lobby for the changes. They discussed the studies in meetings, used the studies to answer questions they raised earlier, and prepared cover letters for transmitting the reports to other groups.

The local officials used the information to:

- alter programs they managed;
- develop grant applications to remedy identified problems; and
- request changes by other agencies with which they worked.

The officials claimed that the studies identified gaps or problems in juvenile services in a manner that permitted them to take remedial actions. The use occurred by direct provision of results to them by the evaluators, by meetings with the evaluators and Committee to discuss the results, and by receipt of the final reports. Three of the officials cited the existence of the information produced by someone outside their organizations as being extremely useful. They were aware of the problems documented, but had been unable to convince others of them prior to the evaluation.

(2) Users Cited Instances of Using the Juvenile Studies: Several instances were cited of where the utilization of evaluation information could be associated with decisions and with changes in the local criminal justice system. The uses cited fall into three categories:

- Informative

  As cited above, the evaluation studies were a source of information for participants in public meetings and hearings concerning organization of services and location of jail facilities.

- Administrative and Program Changes

  Several new or modified services were introduced by juvenile justice agencies to overcome problem areas or service gaps identified by the studies (e.g., nurse practitioners added to Juvenile Hall to increase health services; Recreation Coordinator added to Probation Department; intake procedures modified to better serve short- and long-term detainees; and changes in group counselor classification).

Modification of juvenile recordkeeping procedures by local communities to make them more uniform throughout the County.

- Development of Grant Applications

  Information from the studies was used to support two grant applications aimed at resolving juvenile justice problems identified through the studies.

(3) There Is Evidence that the Use Cited Was Necessary to the Performance of Important Functions: The effort put into the evaluation efforts by the users demonstrate the value they attributed to the activity. The Juvenile Justice Committee members met routinely, reviewed products and the progress of the work, helped obtain data, and actively participated in the dissemination of results.

The availability of the evaluation information was cited by several persons interviewed as being essential to initiating action on problems. Agencies were aware of the problems, but required documentation of the issue to convince others to take action.

(4) Evaluations Used Were Attributable to the MEP: Some evaluation activities of the type supported were planned and partially initiated prior to receipt of the MEP. The MEP assumed and expanded these activities. Without the MEP, the number of issues examined would not have been undertaken.

e. Jacksonville

The Jacksonville RPU utilized the MEP to expand an existing Evaluation Unit to provide additional planning and evaluation information to local decision makers. The pre-MEP Evaluation Unit focused on the evaluation of LEAA funded projects; the MEP unit expanded study areas to cover other parts of the criminal justice system.

The seven studies were found to be useful by local officials, and some use could be cited. Although many program decisions were associated with the studies, it is unclear to what extent those decisions were affected by the studies.

(1) Local Officials Claim Evaluation Information Useful: Eleven local officials were interviewed in Jacksonville to determine whether or not the studies conducted under the MEP were useful to them.[1] Three of the interviews were conducted prior to the end of the MEP, but after completion of studies which were of specific interest to the persons interviewed. The other eight interviews were conducted after the MEP grant had terminated. Most of the persons interviewed were satisfied with the studies of programs on criminal justice components with which they were familiar. Sample responses were:

> Had positive feelings about what they read; e.g., . . . is one of the most significant studies in a long time, and good product.
>
> Felt that what had been read had been useful, e.g., reports reinforced some ideas already held; report provided clout to implement policies and procedures which might otherwise have been resented; some planned to use reports in future decision making; committee formed to decide how to implement recommendations of the Correctional Master Plan.

---

[1] Those interviewed included: two judges; the sheriff; a member of the sheriff's staff; a court administrator; the Director of Jacksonville's Department of Human Resources; the Director of Jacksonville's Department of Health, Welfare and Bio-Environmental Services; Chief Administrative Officer of the City of Jacksonville; the Director of an evaluated program; and the Director of the Community Relations Commission.

All responses were not positive, however. One person expressed neither positive nor negative feelings, and another negative feelings.

> Recommendations . . . not applicable
> . . . and probably illegal.

The Correctional Master Plan appeared to be the most widely read and most controversial of the studies. It was viewed by several officials as containing recommendations worth implementing or worth further study and, in fact, the MEP Unit was asked to do a cost analysis for one of the programs proposed in the study.

(2) Some Uses Cited: Some use could be cited from six of the studies. Most activity was associated with the Correctional Master Plan Study:

> Criminal justice agency personnel and City Council members made decisions based, in part, on information produced, e.g., to build a juvenile shelter, postponement of decision to build a jail, introduction and partial passage of legislation which would enable Jacksonville to form a Department of Corrections independent of the sheriff. Initiated a pre-trial intervention project.

Other uses cited were:

- included evaluated project in the RPU's annual funding plan;
- changes in project activities;
- discontinuation of project which had uncertain future before the study.

(3) Some Information Necessary to Performance of Important Functions: The Correctional Master Plan Study does seem to have been critical to the planning process. The information was used to affect decisions in a process that was initiated independent of the MEP. The need for and importance of the evaluation information is less evident

in the other studies. However, six of the seven studies performed were requested either by a local criminal justice agency or the advisory council to the RPU. These requests indicate that at least at one point in time, the information was considered as needed by local decision makers.

(4) Information Use Attributable to the MEP Project: The major uses that did occur can be attributable to a combination of existing issues and the provision of appropriate information on the issues. While the MEP did not create the issues in all cases, a component of the Jacksonville system involved seeking out and responding to issues. Thus, the provision of appropriate information in a manner that increased the chances of it being used can be attributed to the Jacksonville evaluation system.

f. Association of Central Oklahoma Governments

The Central Oklahoma MEP produced six demonstration evaluations aimed at providing useful information to the management of the evaluated projects. While those receiving the information claimed it was useful, they could cite few uses. Further, the uses cited did not demonstrate that the evaluation information supported a critical function. On the whole, the Central Oklahoma approach did not create or demonstrate the existence of an evaluation market, or show that evaluation made an important difference in local criminal justice operations.

(1) Local Officials Claim Evaluation Information Useful: Seven local officials were interviewed concerning the utility of the evaluation information produced by the Central Oklahoma MEP. The officials were the managers of the projects that had been evaluated, and were in a position to be aware of all local uses of the MEP evaluations. All seven were aware of

the contents of the evaluation of their respective projects. They had read the reports, discussed them with the Central Oklahoma staff, and disseminated them to others.

All the local officials believed the reports had been of benefit to them, and the agencies in which they worked:

- Gave an understanding of what was happening that was previously unavailable.
- Indicated the need for better data collection.
- Documented some problems.
- Provided information that would support funding requests.

Only one person indicated how the studies would have been more useful by stating they were too descriptive and did not identify preferable service alternatives available in the projects.

Five of the officials claimed that the evaluations were of sufficient use that they would participate in them again if given the opportunity. Also, five of the officials claimed that, given the choice between applying resources to evaluation or to additional project activities, they would chose evaluation.

(2) Few Uses Cited: While all projects claimed the evaluation process and products were useful, few actual uses that affected criminal justice system operations could be cited. The uses cited by officials interviewed and in Central Oklahoma documents were limited to:

- modification of record systems (four projects);
- support grant funding requests (two projects);
- modify project design (one project).

The first use is actually due to the evaluation process, as well as the information produced. While conducting the evaluations, Central

Oklahoma personnel encountered discrepancies and gaps in existing record systems that inhibited the evaluators. Modifications were developed in these systems to meet the needs of the evaluators and project managers. Those modifications were continued by the local agencies after completion of the studies.

Information from two of the evaluations was used in applications for continuation grants. The grantees claimed the information was requested by the Oklahoma Crime Commission, but they do not believe that it affected funding decisions. If the information had not existed, funds would still have been granted. Central Oklahoma evaluators shared this opinion.

One project manager identified a part of his project that was modified on the basis of evaluation information. The project included a component aimed at increasing public awareness of the project activities. The awareness was discovered not to exist, and remedial actions were proposed.

(3) Use Did Not Serve a Critical Function: On the basis of the following factors, it is concluded that evaluation did not serve a critical function: level of local resources committed; uses cited; and lack of an existing decision process to be served.

The local projects cooperated in the evaluations, but expended little staff time and resources on them. A decision was made during the first month to abandon plans to ask the projects to collect new data for the evaluations. The RPU management concluded that it would be unreasonable to levy data requests when grant funds had not been made available to the projects to provide data, or the request had not been made prior to award of the project grant. The projects cooperated with the evaluations by: reviewing and approving evaluation designs; submitting to interviews; making project and agency records available to the evaluators; and by reviewing and commenting on final products. The cooperation was judged excellent by both evaluators and managers, and no major problems with the relationship were identified by the managers. However, major parts of the planned analyses had to be dropped due to the lack of data nd restrictions adopted on requesting the projects to produce additional ata.

As noted above, there were no uses cited that affected the criminal justice system's operations in a significant way. A review of existing record systems could have produced the majority of uses achieved. Given this level of use, it was concluded that a real need for evaluation information was not discovered or demonstrated.

The primary customer for evaluation information was the management of the project or the agency operating the project. The evaluation results were given to the managers in draft form and as final products. The managers were free to utilize the information as they saw fit. Two of the six evaluated projects incorporated the evaluation into the process of requesting continuation of the projects, and all claimed to have disseminated copies to local officials having contact with the projects. But none of the managers interviewed could identify a local or state decision process that depended on the information to function.

(4) Evaluations Used Are Attributable to the MEP: The evaluations were completely funded with MEP funds and would not have occurred

without the MEP. Other evaluations were conducted concurrently on half of the projects as a result of grant requirements imposed by the Oklahoma Crime Commission. Thus, some evaluations would have occurred, but not of the type undertaken by Central Oklahoma.

5. KNOWLEDGE DEVELOPMENT

No site was able to develop information to test in the Knowledge Development market. Three sites had planned to do so, but fell short in practice. Two different strategies were attempted--one relying on standardized project data reported by subgrantees, the other intensive multi-project evaluations of a program area. The results of these efforts raise questions about the feasibility of the "knowledge" building evaluation.

Two sites attempted to rely on reported data, Illinois and Michigan ("Standard"). The Illinois system originally was to link project data with data on area crime and population characteristics to carry out computerized program impact analyses. The system was never implemented because of several operational problems. For example:

- The standard designs for a program area frequently fit only one project.
- The link between computerized project and socioeconomic data bases was never set up.

An Illinois senior manager summarized their experience as follows:

> When we started we were shooting for the moon, but we didn't have a rocket or know how to build one. Now our sights are set lower, but we know how to get there.

The Michigan standard reporting system attempted to measure program process and outcome indicators; it did not attempt cost-effectiveness and impact analyses. In addition to being a monitoring tool (project oversight), Michigan hoped the standard system would enable it to draw conclusions about the effectiveness of program concepts. Implementing the system required some accomodation of the standard instruments for variation among projects in a program area. At the time this report was written, Michigan had not completed any cross-project analyses. There is, however, a question about the validity of the standard instruments, and whether the project data can be aggregatd for a program area. This issue is raised dramatically by the two contractor intensive evaluations carried out under the Michigan MEP.

There were two attempts at doing intensive evaluations of program concepts: Michigan's contractor evaluations and Pennsylvania's phased approach to evaluation. Michigan carried out intensive evaluations of two program areas which indicated that the programs had no effect on crime, criminal justice system or clients. The contractor labeled this an implementation failure, rather than a failure of the concepts tested. He concluded that the projects' operations and objectives deviated markedly from what was described in plans and subgrants, and expected by the SPA. The concepts to be tested had never been implemented.

Pennsylvania tried a phased approach to evaluation--the first phase was to group similar projects, develop measures, and identify "best" project models; the second phase would be a more intensive evaluation of project models. Pennsylvania was unable to complete Phase I. One of the reasons cited was the great diversity of projects funded under any one program element. Common objectives and measures were not readily apparent.

Evidently, two characteristics of SPA/RPU projects tend to disable the "knowledge" building evaluations:

- Variation among Projects within Program Areas

  Some agencies retrospectively attempt to group projects according to similar activities and/or similar objectives. They find that projects with similar names (i.e., Youth Service Bureaus) or projects funded under a common program title (i.e., Youth Diversion Projects) have major differences in activities and purpose. For evaluation purposes they cannot be grouped and characterized by similar measures.

- Difference between Project Operations and Agency Rhetoric

  A problem that can occur with or without variation among projects is that the projects are doing drastically different things than assumed by the agency. Thus, the objectives and evaluation questions the agencies have are not appropriate. An evaluation will show "no effect" in terms of agency objectives because the projects are not doing what the agency thought they were doing.

These characteristics can be traced to the inability of agency proposal review, technical assistance, and monitoring functions to assure the level of project definition and implementation needed to do "knowledge development" evaluation.

E. SITE CONTINUATION OF THE EVALUATION SYSTEM TESTED

The ultimate test of the utility of the evaluation system is whether or not it is continued. Of the eight completed MEP sites, three plan to continue a version of the systems they tested: Jacksonville, Ventura and Michigan. However, even in these three cases, the systems tested will be significantly modified. Table 3 summarizes the continuation results.

In dropping the MEP evaluation systems, the sites went to some alternative; in most cases, to what they had before MEP, in a few cases to a

TABLE 3: CONTINUATION OF THE TESTED SYSTEM

	Continuation Decision on Tested System	Site Rationale for Decision
Michigan (Reporting System)	Modified and Continued	System essential to SPA subgrant management strategy.
Ventura	Modified and Continued	High demand for technical assistance from local agencies.
Jacksonville	Modified and Continued	Demand for studies was confirmed.
Michigan (Intensive)	Not Continued	Contractor--SPA interactive design process was disruptive. Intensive evaluations will be performed under a different process.
Illinois	Not Continued	Reporting system too expensive to implement across all projects.
Alameda	Not Continued	Studies controversial and evaluation capability too expensive. In-house staff vulnerable to diversion to non-evaluation tasks.
Central Oklahoma	Not Continued	Funding not available
Columbia	Not Continued	Funding not available
Pennsylvania	Not Continued	Original Phase I-II methods could not be implemented.

new approach. Sometimes it is arguable whether a system was being continued or discontinued. Several sites claimed they were continuing the system when, in fact, no essential components of the system remained. In this report a site is said to continue its MEP evaluation system when it has formally allocated resources to: (1) carry out activities tested under the MEP; (2) produce products whose utility was demonstrated under the MEP; and (3) carry out those activities tested that are essential to produce the information product. Sites can modify systems by deleting unnecessary activities, and by changing who carries out the tested activity--and are still viewed as continuing the system tested.

1. CONTINUATION: MICHIGAN, JACKSONVILLE, AND VENTURA

a. Michigan (Reporting System)

Michigan pilot tested a reporting system on four program areas. The evaluation system had the following characteristics:

- a design phase carried out by SPA program staff;
- standardized instruments used by subgrantees to collect and report data;
- interns collecting data on site and monitoring the reported data;
- central, computer data processing;
- project tabulations and multi-project (program) analyses of process and outcome variables;
- a new evaluation office with formal responsibility for managing the evaluation system.

Continuation would mean expansion of the reporting system to other program areas, and retention of sufficient system components to assure production of useful information.

Funding limitations have caused Michigan to modify the data collection, data processing and data analysis components substantially. Three of the four intern positions have been dropped, thus limiting the on-site collection and monitoring of reported data. Some of these responsibilities will be transferred to SPA program staff. For new program areas, with the exception of juvenile projects, there will be no central, computerized data processing; subgrantees will submit summary data.

Despite these changes, it appears that Michigan will be able to produce the project tabulations and multi-project comparisons, while retaining several of the tested features that made the system work: the design phase, standardized instruments, and an SPA evaluation office with responsibility for the system. The SPA has established a two person evaluation office (plus one intern), and already has program staff carrying out design work for additional program areas under the direction of the evaluation office.

There is some question as to whether the tested system will be successful as modified. The tested system had student interns carrying out a number of functions (site data collection, monitoring data flow, data processing, report writing) which now will have to be carried out by the SPA program staff. It remains to be seen whether the program staff have, or can create, the time needed to do the work. Michigan's policy is that they will.

b. Ventura

The system tested by Ventura had the RPU evaluation staff (one person) serving as an evaluation resource for local criminal justice agency personnel. The experience indicated that this resource was very much in demand by local agencies.

The system tested had the following characteristics:

- a demand for evaluation assistance from local agencies;
- an RPU evaluator with evaluation skills and resources available to use in technical assistance;
- an RPU evaluator with the ability to diagnose the information needs of local officials and meet them.

Continuation would require the RPU to secure funding for the position and fill it with the appropriate person.

Ventura plans to continue to provide the service but at a lower level of effort, due to funding constraints. A position for an evaluator was funded through a subgrant, and at least half the evaluator's time was to go to the technical assistance function. In addition, funds from the Comprehensive Employment and Training Act are being used to create additional evaluation positions which the RPU's evaluator will be able to use.

The uncertainty in continuation lies in the recent departure of the MEP evaluator. Since local observers attributed much of the success of the MEP to this person's technical and interpersonal skills, it is not clear whether the RPU will be able to find a replacement. However, the RPU still plans to "provide technical assistance to any criminal justice agency undertaking evaluation studies." The position has been created and recruitment is underway.

c. Jacksonville

Continuation of the Jacksonville Project would involve maintaining the in-house evaluation staff and performing evaluations in response to local requests. The in-house staff has been maintained at a lower level (four) than during the MEP (six), but at a higher level than before the MEP (two). The remaining staff is performing studies similar to those conducted during the MEP, and responding to local requests for evaluation information.

The size of the in-house staff is a function of available funds, experiences with the level of demand, and prior problems with increasing staff size. During the MEP, demand for information did not exceed the staff capacity. Studies were performed in response to local officials' request for information. Based on the requests received during the MEP, the RPU believed that requests could be met with a slightly smaller staff than maintained during the MEP.

Upon receipt of the MEP, the RPU doubled the size of the existing evaluation staff. Also, several times during the grant period, staff turnover necessitated recruitment and training of new personnel. These staffing activities were found to be time consuming and, on several occasions, did not result in the needed evaluation capability. Consequently, when through attrition the MEP grant ended with a competent but smaller evaluation staff, the RPU was content to stay at that level and avoid capability building problems encountered during the MEP.

2. DISCONTINUED

Five sites did not continue the MEP systems tested: Michigan (Intensive); Central Oklahoma; Illinois; Columbia; Alameda; and Pennsylvania. In four sites discontinuation is clear cut and acknowledged by the site. In two cases the conclusion is less clear cut--Michigan (Intensive), and Illinois. Here some part or version of the tested system was carried on.

However, we concluded that the change in the product or the absence of system components necessary to produce information were so severe that one could not say that the tested system was judged successful and is being continued.

a. Michigan (Intensive)

Michigan pilot tested multi-project impact and effectiveness evaluations on two program areas. The evaluation system had the following characteristics:

- one large contract to a university research center;
- a phased contract calling for a design phase in which Michigan's information requirements would be documented and analyzed;
- program specialists guiding the contractor design work.

Continuation of this component would mean that Michigan would carry out several intensive evaluations of programs annually using a university research center and the phased contract approach.

Michigan has decided to do three to four intensive evaluations a year. They will be contractor efforts and funded with grant monies, rather than through an SPA budget item. The contractor will receive guidance in a design phase from SPA program staff, and will be monitored by and responsible to the SPA evaluation office. However, it is concluded that the tested system is not being continued for two reasons: discontinuation of the phased approach and the feasibility issue raised by the field test.

Michigan is dropping one design activity from its tested system. The contractor will not be asked to collect data to identify SPA information needs. This activity was viewed as disruptive and non-productive by the SPA. Instead, the contractor will develop a design based on discussions with program specialists and subgrantees, and submit the design to the evaluation office for comment.

It was, however, this interaction among contractor, program specialists and management that made the intensive evaluation system different than what Michigan had done in the past. Many agencies hire contractors to evaluate a number of projects. The design phase in the MEP field test raised issues, affected what the contractor did, and affected the users.

Finally, as a result of the field test, there remains some uncertainty about the feasibility of producing information on what effects projects have under different conditions from intensive evaluations. Recent intensive evaluations have found no effects of programs, and have raised issues about whether programs and projects as designed and implemented by the SPA have plausible relationships between activities and effects which could be tested. The SPA has decided to place more staff time in giving assistance to subgrantees during project design and implementation phases. Whether the SPA can successfully implement a number of similar projects with similar objectives remains to be seen.

b. Illinois

Illinois used its MEP to develop standardized reporting instruments that would cover all evaluable funded projects originally estimated to be about 80 percent of those funded. The evaluation system to be set up had the following characteristics:

- project reporting instruments standardized for each program area;
- all evaluable programs covered;

- central computerized processing and analysis;
- routine project and multi-project analyses of project processes, outcomes and impacts.

Continuation would require establishment of the data flow from subgrantees, and funding of the central processing and analysis capability in the SPA.

Illinois' experience with the system was negative--it was not able to implement across all eligible projects the data flows and data processing capabilities it had envisioned. During the MEP, Illinois had been able to obtain and process data from only 25 percent of the projects provided data collection instruments. Moreover, in most cases, the standardized instruments could only be applied to one project in one class, making multi-project analyses infeasible.

As a result of these problems, a decision was made to drastically curtail the system rather than invest the resources and energy required to implement it fully. Illinois planned to continue to use the standard instruments on about one-fourth of its projects as a monitoring tool. There are no plans to perform on a continuing basis the project and multi-project outcome and impact analyses using standardized forms.

c. Central Oklahoma

Central Oklahoma tested a project evaluation system in which RPU evaluation teams used readily available project data augmented by selected surveys to do quick evaluations of project performance. The evaluation system had the following characteristics:

- routine short duration project evaluations;
- a standard evaluation framework followed on each study;
- RPU in-house evaluation teams carried out data collection and analysis.

Continuation would have required funding of the RPU evaluation positions.

Central Oklahoma believed its system had sufficient merit to continue but, given other demands, could not fund these positions with Part B allocations. Attempts were made to secure funds from other sources but they were not successful.

d. Pennsylvania

Pennsylvania attempted a phased approach to intensive evaluation. Phase I evaluations would rely on past evaluations and research to develop project models, evaluation designs and estimated performance levels. Subgrants would be let to implement the models. Phase II evaluations would be carried out on sets of projects funded by Pennsylvania, and use quasi-experimental designs to improve the knowledge base. Pennsylvania found that the research literature, grant files, and project evaluation reports did not provide sufficient data or definition to support its Phase I-Phase II design strategy. Prior to the end of the MEP, the old strategy was abandoned.

e. Columbia

The Columbia MEP Project primarily consisted of the temporary creation of a regional evaluation resource person. Continuation of the project would have meant continuation of the position, but it did not occur. Some of the data collection and analysis procedures introduced to local agencies are being continued by those agencies, but the RPU role has been dropped.

The RPU did not have funds available to continue the position, and sufficient interest did not exist to obtain the funds by eliminating other RPU activities or seeking funding from other sources. As the MEP grant entered the last few months, the RPU decided not to attempt to secure funds for the position, and the evaluator began to search for other positions. One month prior to grant termination, the evaluator accepted other employment, and the position was not filled for the remaining grant period.

f. Alameda

The Alameda MEP Project consisted of the creation of in-house study teams to conduct sizeable evaluations on major issues or segments of the criminal justice system. Continuation of the approach tested would be continued funding of the in-house study team, and the undertaking of similar evaluations. This has not occurred.

Financial, feasibility and practicality reasons have been cited for not continuing the MEP activities. While the MEP was underway, the RPU experienced significant budget cuts that necessitated a nearly 50 percent staff reduction, and reduced the possibility of continuing the tested type of evaluation at the same funding level. However, several members of the local evaluation and policy-making community believe that the MEP effort demonstrates that the feasibility of performing impact evaluations is low, and the risk of failure is too great to warrant the expenditure. Questions were also raised by members of the RPU staff and policy-making Board as to whether it was practical for the RPU to conduct studies covering activities not funded by LEAA. Several local officials stated that there might be a conflict in the role of evaluating local criminal justice agencies' activities, and having representatives of the agencies on the Policy Board of the RPU. However, the situation was discounted by the RPU Director and several Board members. Yet, the evaluations clearly caused some resentment in the agencies whose representatives are on the RPU Policy Board.

The Alameda County government established an Evaluation Unit the same year as the MEP work was initiated. Initially, the County Unit did not perform evaluations in the criminal justice area, but does now. Also, the Alameda RPU awarded grants to the County government to conduct sizeable evaluations. Thus, while the approach to evaluation tested during the MEP has been discontinued, major criminal justice evaluations are still being funded within Alameda. The trend is not attributable to the MEP, however, but appears to be due to the continuing local demand for evaluation information.

F. FACTORS WHICH AFFECTED SUCCESS

The two factors on which MEP success was judged are continuation and usefulness of information. While no complete set of neccessary or sufficient conditions for success was identified, certain factors which contributed to inhibited success were identified.

1. SUCCESS IN HAVING TESTED SYSTEM CONTINUED

Two conditions appear to affect the continuation decisions of the MEP sites once the production of evaluation information is shown to be feasible: (1) the significance of the evaluation system to management's agenda; and (2) availability of funds.

The three sites that continued the systems had very visible management activities which the tested evaluation systems supported. Michigan SPA management had a strong role in subgrant planning and funding decisions. Its policies emphasized staff working with grantees and providing them assistance. Michigan's reporting system, stressing heavy training and participation of the program specialists, was a natural supporting function for the management strategy.

Ventura's activities included an effort to get local agencies to adopt goals and standards, and to evaluate their programs with respect to these goals and standards. The Ventura RPU believed that local agencies would have to develop evaluation capability and an inclination toward using evaluation. A similar situation existed in Jacksonville. Since the Jacksonville RPU is a division of the Mayor's office, its agenda included providing assistance and information to the Mayor, City Council, and City agencies. The MEP enabled the Jacksonville RPU to respond to local requests for information. The result of the MEP encouraged the RPU to continue offering technical assistance.

In the other sites it was not always as clear cut that the evaluation system was supporting major agency initiatives. During the period when the MEPs were ending, most SPAs and RPUs were facing funding cuts. Many of the sites were experiencing or expecting reductions in "Part B (Administrative) Funds" as the MEP grants ended. Therefore, the funds that might have been available to continue the MEP were shrinking. This presents particularly severe funding problems for RPUs.

Table 4 shows the MEP grants relative to the size of the SPA or RPU. In the case of the RPUs, the MEP grants were as big as the whole RPU administrative budget (Part B funds). Therefore, at least for the RPUs, continuation would not have been a straightforward matter, even under

TABLE 4: BUDGETS AND STAFFS OF MODEL EVALUATION PROGRAM GRANTS

	Estimated Size of SPA or RPU (1976)		
	Administrative Budget[1]	Professional Staff[2]	Size of MEP Grants[3]
Central Oklahoma	$ 61,000	3	$108,000 (one year)
Illinois	$2,773,000	58	$250,000 (one year)
Pennsylvania	$2,930,000	59	$261,000 (two years)
Ventura	$ 122,000	2	$ 74,000 (one year)
Alameda	$ 333,000	15	$144,000 (one year)
Michigan	$2,280,000	45	$248,000 (two years)
Jacksonville	$ 109,000	15	$ 85,000 (one year)
Columbia	$ 59,060	3	$ 42,000 (two years)

---

[1] Administrative Budget, primarily LEAA Part B funds.

[2] Professional staff are positions authorized for criminal justice activities.

[3] The two year MEP required a 50 percent match in the second year. The figures given here are the original grant sizes.

the best of conditions. As allowed by LEAA, two MEP sites continued their evaluation programs by building evaluation components into subgrants (Part C funds) in both their administrative budgets and subgrant budgets. At least four of the sites (RPUs) indicated that securing funds was a problem and a major reason for not continuing the MEP.

2. SUCCESS IN PRODUCING USEFUL INFORMATION

Four sites were singled out as being successful in producing useful information. When they are contrasted with the less successful sites, two conditions are found to be associated with success. At the successful sites there was one or both of the following conditions: (1) a demand for evaluation information already existed; and (2) the users felt they owned the study.

As described earlier, in Ventura, Michigan and Alameda (County Board studies), and Jacksonville, there is evidence that the MEP was attempting to satisfy a demand for evaluation that already existed. In these cases the users' interest was the reason for doing the evaluation. Michigan's management policies required evaluative information; in Ventura local agencies were undertaking studies on their own initiative; in Alameda (County Board component) information was required to estimate the number and size of new jails the County had to build; and Jacksonville officials identified several areas of information need, including information on detention operations.

Several less successful sites attempted to create a demand, expecting the evaluation product to sell itself. In Central Oklahoma and Illinois, the users, selected project directors in one case and SPA program staff in the other, expressed little interest in the evaluation effort before or during the MEP. The final MEP product did not stir either set of users to significant action.

Pennsylvania hoped to create users as well as user interest. The SPA did not have the planning functions in place, which the Pennsylvania MEP was designed to support. One expectation of the Pennsylvania MEP unit was to use the evaluation information to convince the SPA that it should set up these functions. Paradoxically, the absence of the proposed SPA planning functions contributed to the MEP bogging down and failing to produce information.

The message appears to be that one's chances of success are much higher if one identifies an existing demand, rather than try to create the demand.[1] However, the second condition, user ownership of the study, may be a way to go about creating a demand.

The three successful sites all operated in a manner that allowed users to have ownership. That is, the users saw the study as being done for them, asking the questions they were interested in, and answering them in a way that they understood. Ventura often waited until the agency initiated the evaluation effort and then gave assistance. Selection and definition of the evaluation was up to the user. In Michigan the program specialists designed the standard evaluations themselves. Michigan's intensive evaluation contractor documented and fed back management's perception of the programs and information required. Alameda put the evaluation teams under the administrative control of the user group.

---

[1] This is similar to the conclusion reached by a two year Office of Economic Opportunity funded experiment in the Atlanta school system. An evaluation system was designed and implemented in Atlanta to provide information on school performance which management said it needed. A follow-up study showed that the school system did not use the information, primarily because there were no administrative and management functions designed to use it. Simply providing information is not likely to lead to action. (See Bayla White et al. The Atlanta Project: How One Large School System Responded to Performance Information, Paper 9-0507-5, Washington, D.C., The Urban Institute, March 1974.)

One part of the Alameda MEP represents a situation where a need for evaluation may not have pre-existed. Alameda's studies of juvenile justice were done under the direction of a committee representing local agencies, advocacy groups and the public at large. The committee was formed to direct the studies, and took responsibility for delivering the product. The committee invested a great deal of time in exercising their responsibilities, and became advocates for the changes suggested by the studies. In a sense, this part of Alameda's MEP did create a user group and user interest. The mechanism was to give ownership of the study to the potential user. Ownership does not seem to occur when the evaluator decides on the design and merely submits it to the user for review and approval.

As discussed in the next chapter, several of the MEP sites included activities in the development and operation of evaluation systems that were designed to take advantage of existing demand for evaluation information, or involve users in evaluation operations. The results of those efforts generally increased the utilization of the evaluation produced.

III. BUILDING AN EVALUATION CAPABILITY

The Model Evaluation Program was part of an LEAA effort to build evaluation capability within the criminal justice community. This chapter addresses the strategies the MEP grantees utilized to develop evaluation capabilities, and presents the types of capability they attempted to develop, and those that were developed. Section A summarizes the experiences of the sites and emphasizes the fact that all sites except one were successful in establishing some new evaluation capabilities and producing evaluation information. Section B describes the strategies used to develop evaluation capabilities, and the changes in strategies and expectations that were made. Section C addresses cost issues, and Section D presents a synthesis of the characteristics of the various evaluation systems developed. The chapter concludes with brief descriptions of the eight sites that completed their MEP grant during 1977.

A. PLANNING AGENCIES WERE ABLE TO ESTABLISH EVALUATION SYSTEMS

The Model Evaluation Program demonstrated that by earmarking funds for evaluation, SPAs and RPUs are almost certain to set up an internal evaluation capability which produces or helps produce information. The eight completed sites also demonstrated that the SPAs and RPUs are unlikely to be successful in establishing an evaluation capability within local criminal justice agencies.

Seven of the eight completed sites were successful in establishing an evaluation system, though not always the one they had originally planned. The MEP demonstrates that by establishing an organizational unit with resources and responsibility for evaluation, an evaluation capability can readily be established. However, as discussed in Chapter II, its utility and continued support is another matter.

Originally three RPU MEPs expected to establish evaluation systems or capabilities in local criminal justice agencies. None were able to implement a strategy for doing so. Consequently a local agency evaluation system established by and running independent of the RPU was never tested. Each site fell back to testing some evaluation capability in the RPU. In the end, all MEP tested evaluation systems were those in which the evaluation capability rested within the SPA or RPU.

The evaluation systems tested varied considerably on a number of operating characteristics: study methods; production levels; types of SPA/RPU evaluation product; cost per product; standardization of the production process; and mechanisms to involve the user. The only system characteristic associated with success on the use criteria was the involvement of the user (or "ownership" as discussed in Chapter II). The variation in evaluation systems reflects the wide variation in how State Planning Agencies and Regional Planning Units perceive their mission and function. Under the current federal regulations, there is no "right" or "wrong" evaluation system.

The Model Evaluation Program experiences offer some guidance on the cost of developing and operating an evaluation system. Some systems have high development and low operating costs, while for others the reverse is true. The amount to invest in any given system seems flexible--one can buy a big version or a little version. There is no information available to show the difference investment in evaluation makes to SPA and RPU performance. It is important to note, however, that a "useful" system has hidden costs which may not show up in a typical evaluation ofice budget. These are usually costs borne by groups outside the evaluation office.

Once funds are earmarked for evaluation, implementation of an evaluation system appears to be enhanced when there is active support from SPA or RPU management, and/or a commitment to produce products for specific people. The only site where neither of these conditions were met failed to implement a system.

This chapter describes the evaluation systems tested and how they were established. The strategies and costs involved in setting the systems up and testing them are identified first, followed by descriptions of the systems.

B. EXPECTATIONS AND STRATEGIES OFTEN HAD TO BE CHANGED

Each MEP site started with a vision of an evaluation system and a strategy to implement the proposed system. In five out of eight cases, as the site gained experience, the expectations and the original strategy had to be significantly changed. Of the eight sites completed, three sites carried out their original strategy, four sites changed their strategy and carried out a new one, and one site failed to carry out any strategy during the grant period. The changes were required because in planning their Model Evaluation Program, sites had either misjudged the market for evaluation or had overestimated the feasibility of producing a type of information. The changes included redefining (1) where the evaluation system was to be set up, (2) the audience for the information, (3) the type of information to be produced, and (4) the implementation strategy. When establishing an evaluation capability, one should be prepared to shift expectations and strategy guided by initial experiences.

1. CHANGING EXPECTATIONS AND STRATEGIES

Table 5 summarizes the degree to which sites had to change their expectations about the kind of evaluation system that could be established.

a. Sites Where the Original Expectations Were Met

The three sites able to carry through their original strategies were Michigan, Alameda and Jacksonville. Each site was able to deal with problems that threatened implementation.

Michigan wanted to institutionalize both the design and use of evaluations as functions and responsibilities of the SPA management and program staff. Michigan's strategy was: (1) to increase the evaluation capability of the whole SPA staff through training; and (2) to pilot test two evaluation systems in which management and staff heavily participated. The strategy involved a large training program involving all staff, development of prescriptive material to guide the design and review of evaluations, and pilot testing of a formal design and implementation process.

Several factors threatened the implementation of the strategy, notably early termination of the training contract and turnover in the evaluation coordinator position. However, the MEP project was able to deal

TABLE 5: SITES THAT MET OR CHANGED EXPECTATIONS

MEP Site	Original Expectations Met	Changed Expectations Met	No Expectations Met
Michigan	X		
Alameda	X		
Jacksonville	X		
Ventura		X	
Central Oklahoma		X	
Illinois		X	
Columbia		X	
Pennsylvania			X

with these problems and continue. One significant problem fell outside the control of the MEP. During the training course, staff attendance and participation began to fall off. Michigan's top management intervened to support the MEP, and attendance improved. It was made clear that management was attempting to institutionalize evaluation, and everyone was expected to participate.

Alameda wanted to demonstrate that evaluations which compared innovative projects with more traditional programs would be feasible and useful to local decision makers. The system would consist of RPU staff doing studies under the direction and guidance of local decision makers. Since the primary feature of the system was an organizational one, Alameda's strategy was to acquire the staff and pilot test the system. There were no developmental or training steps.

The strategy did not run into any serious obstacles. The major problem encountered was the expansion in the scope of work and the delays caused by the involvement of the users.

Jacksonville wanted to expand its evaluation unit in order to build evaluation capability in Jacksonville more rapidly than previously possible, and have a greater effect on the local criminal justice system. It was by participating in the MEP that Jacksonville hoped to build evaluation capabilities in two ways:

- conduct evaluation studies which would be so useful to local criminal justice decision makers that funds would be allocated to continue the evaluation unit after the MEP; and
- assist programs or projects within the Jacksonville criminal justice system in developing or upgrading existing in-house data collection and evaluation procedures and capabilities.

The RPU did conduct a series of studies which partially achieved the first objective and did have effects on the local criminal justice system. The studies were a bit longer than originally envisioned, and seven instead of the planned ten were performed, but they served the intended purpose.

A mechanism was not developed to directly assist local agencies up-grade existing capabilities, and the objective was abandoned. The RPU focused efforts on developing and demonstrating its own evaluation capability.

b. Site Where New Expectations Were Met

Four sites had to change their strategy and their expectations during the grant period, but were able to test a system: Ventura, Central Oklahoma, Columbia and Illinois. The first three had hoped to set up evaluation capabilities in local criminal justice agencies. These strategies did not work out, and they fell back to implementing and testing an RPU evaluation system. The fourth site, Illinois, found that the system it had hoped to put in place was not feasible, and so it reduced the scope and types of evaluation it planned to produce.

The Ventura RPU had expected to establish self-sufficient evaluation programs in local agencies, an event which would have complemented the RPU's effort to have local agencies adopt the "goals and standards" of the National Advisory Commission on Criminal Justice Standards. Ventura planned to use a combination of classroom training, on-the-job training and demonstration of evaluation utility to build this evaluation capability in local agencies. Early in this grant period, the RPU concluded that the training was not having the payoff expected. Since there was a demand for evaluation assistance, the RPU changed its strategy to one of pilot testing an RPU technical assistance role.

Central Oklahoma had planned to develop, field test and package an evaluation methodology that would be useful to local criminal justice agencies. The methodology would be transferred to local agencies through training courses. Central Oklahoma was able to carry out part of its strategy--develop, field test and package--but never transferred the method to local agencies. The MEP ran out of funds before the training could occur. Moreover, in the field test, the RPU evaluation staff applied the method, and not local officials. In effect, Central Oklahoma tested an RPU evaluation system in which RPU teams used a standard process to evaluate individual projects.

Illinois envisioned a computerized evaluation system that routinely turned out multi-project outcome and impact studies. The data sources were to be a standardized subgrantee reporting system, and existing crime and socioeconomic data base. The Illinois MEP strategy was to intensify a developmental effort already in progress, and turn out the required standard reporting instruments, geographic data packages and analysis programs. Illinois was able to produce the standard instruments, but ran into major implementation and developmental problems:

- few instruments applied to more than one project, making multi-project evaluation impractical;
- the data flow from subgrantees could only be established on one-third of the funded projects; and
- the geographic data bases could not be set up for impact analysis, as envisioned.

Illinois revised its strategy and tested a project monitoring system using reported data on about a fourth of the SPA projects.

c. Sites Where No Expectations Were Met

Pennsylvania's Model Evaluation Program had planned to produce information that would enable the State Planning Agency to take a strong role in program planning and policy setting. Traditionally, the RPUs carried out these responsibilities, with the State Agency having a very minor role. The Model Evaluation Program planned to develop "optimum models" of program types--descriptions of the program design, program objectives, expected performance and evaluation systems. The models were to be optimum in that past research and evaluation would indicate that they were successful. The SPA would then give implementation and intensive evaluation of the "models" highest priority. Pennsylvania was unable to carry out studies required to produce the models. They found it impossible to extract and synthesize information from the research literature, evaluation reports and the project files. When the original strategy failed, the MEP was faced with no obvious alternative. The State officials had no function which required program evaluations, and the State office itself was in a period of transition and role definition. Attempts to change the MEP expectations and strategy were disabled by the organizational environment in which the SPA Evaluation Unit operated.

2. TYPES OF STRATEGY FOLLOWED

The strategy followed in implementing a system seemed to be less important then the ability to adapt and change when there is evidence that the strategy is not going to work. The strategies finally followed, however, do have implications for costing out a capability building effort, and these will be discussed in Section C below. The different types of strategy followed are identified in Table 6.

The strategies consisted of one or more of the following sets of activities:

- operate the system on a full scale basis to demonstrate it;
- operate the system on a pilot basis to test it;
- develop on a one-time basis components required to operate the evaluation system; and
- provide training in how to do or use evaluation, a one-time cost required to establish the system.

The MEP grant monies went to carry out these types of strategies.

Four of the eight completed MEP sites simply operated full scale the system they had in mind (see Table 6). These systems are more defined by organizational arrangements than by a standard methodology or formal procedures.

Illinois' strategy was to use its MEP resources to develop evaluation designs and computer programs for all its evaluable programs in one major effort. The Illinois system would then operate using the MEP developed products.

Central Oklahoma and Michigan combined several strategies. Central Oklahoma developed a method for doing project evaluations which its study teams used in a full scale demonstration of the system. Michigan developed a method for designing evaluations and standard reporting instruments, and undertook a large training effort involving the whole SPA staff. Michigan then pilot tested both its evaluation systems.

C. COST ITEMS INVOLVED IN SETTING UP AND OPERATING EVALUATION SYSTEMS

The MEP sites provide some indication of the cost involved in setting

TABLE 6: IMPLEMENTATION STRATEGIES FOLLOWED IN SETTING UP THE TESTED SYSTEMS

Sites Completed	Operated Full Scale	Pilot Test	Development	Training
Central Oklahoma	X		X	
Illinois			X	
Michigan		X	X	X
Pennsylvania	X			
Ventura	X			
Alameda	X			
Jacksonville	X			
Columbia	X			
Sites in Progress				
Virginia		X	X	
New Hampshire	X		X	
Washington				X
Massachusetts	X			

up and operating an evaluation system. The demonstration also pointed out:

(1) Given the lack of specific agency objectives and the variation in mission and role definition among agencies, it is impossible to say how big or small an evaluation system (and its budget) should be. The contribution of the evaluation expenditure cannot be associated with some increase in agency productivity or performance. Any of the systems tested could have been bigger or smaller.[1]

(2) In estimating the cost of the evaluation system, one often has to go beyond the cost incurred by an Evaluation Unit. The evaluation system is usually part of a larger management system and, for the evaluation system to work, other pieces of the management system must produce certain products.

(3) Since the products vary in method and use from site to site, one cannot compare the cost per product among sites, although the production levels achieved by each site are known.

Here we describe cost items and how they were incurred in relation to the sites' implementation strategy. Section D describes some of the characteristics of the systems tested.

Most of the MEP budgets went for staff, either in-house or contractor. Expenditures for such items as travel and computer were sometimes significant, but always small relative to personnel expenditures. The major cost items are the expenditures for certain types of staff activity—training, study teams, management. Table 7 lists the cost items for each MEP site.

The costs are broken out to reflect the implementing strategy, and thereby distinguish start-up from operating costs. Table 7 distinguishes agency cost items as (a) training and developmental, or (b) operating costs. The operating costs are further broken down into those incurred by the Evaluation Unit and those by other units. Finally, a third category is identified, hidden cost items, which are essential to the operation of the system, but would not usually show up in an evaluation system budget.

---

1. In a survey conducted in the Fall of 1975, State Planning Agencies estimated that 7 percent of their professional staff time went into evaluation.

TABLE 7: COST ITEMS INCURRED IN SETTING UP AND OPERATING MEP EVALUATION SYSTEMS

Site (MEP Budget)*	Training and Development Items	Planning Agency Operating Cost Items		Hidden Cost
		Evaluation Unit	Non-Evaluation Unit	
Alameda ($144,000)		• Research Coordinator: one-half to three-quarters time • Two research teams: about four to six full-time people, plus some part-time people and interns		User groups: the delegation of responsibility to user groups involves significant cost in user time.
Central Oklahoma ($108,000)	• Research staff: training and method development, and contractor support	• Research staff: three full-time people, plus several part-time interns		
Michigan (Reporting System) ($142,000)*	• Trainer: six man-months • SPA staff: 150 man-days in training (Fundamental and Application groups)	• Evaluation Coordinator: one full-time person • $30,000: travel and data processing • Interns: 90 hours per week	• Program specialists: several man-weeks in design work and introducing instruments to the project	Project reporting cost (never estimated)
Michigan (Intensive Evaluation) ($190,000)*		• Evaluation contract: $190,000 for two program evaluations, about 12 man-years of effort, much of it student. • About 20 percent of contract resources went into a "design" phase.	• Program specialists/management: several man-weeks of staff time on each study providing information, resources and comments.	
Ventura ($74,000)		• Evaluator: one full-time • $40,000: expenditures for consultants and other services		Agencies requesting help bear most of the evaluation costs.

---

*Two year grants include 50 percent second year match.

TABLE 7: COST ITEMS INCURRED IN SETTING UP AND OPERATING MEP EVALUATION SYSTEMS (Cont'd.)

Site (MEP Budget)*	Training and Development Items	Planning Agency Operating Cost Items		Hidden Cost
		Evaluation Unit	Non-Evaluation Unit	
Illinois ($250,000)	Evaluation staff: four full-time people temporarily added for one year to develop all instruments required (augmented existing staff of four)	• Computer analysis: one full-time person programming and analysis • Data entry: two to three part-time people		Project reporting cost (never estimated)
Columbia ($56,000)*	Individuals obtained performed extensive self training	• Evaluator: one full-time person		Projects evaluated performed most of the data collection, analysis and reporting. SPA had an evaluation system in place which the RPU evaluator was attempting to aid.
Jacksonville ($85,000)		• Evaluation staff: staff of two augmented to five full-time people for the MEP		

---

*Two year grants include 50 percent second year match.

Three of the sites had significant start-up costs for method development and training: Michigan, Central Oklahoma, and Illinois. Examination of Michigan and Illinois shows how the implementation strategy affects the evaluation system budget. Illinois used its MEP to develop in one brief time period all the instruments it would require, while instrument development is part of the routine operating cost of Michigan's system.

The operating costs of an evaluation system--those costs encountered in production of information--usually are borne by the Evaluation Unit. However, in Michigan's two systems, other planning agency staff had to be budgeted as part of the system. The two reporting systems--Michigan's and Illinois'--also have cost implications for subgrantees who must collect and submit data. These costs were never estimated on the MEP.

Finally, there are usually other costs which are not really part of the evaluation system or agency budget. They are usually costs that must be borne by the user. These are shown in the last column in Table 7. In Alameda, the user groups convened to administer the evaluations invested a great deal of time. In Michigan, the users had a function that gave them the knowledge required to participate in design work. In Ventura, the agencies requesting assistance had carried out many of the design, data collection and analysis steps with their own resources.

D. EVALUATION SYSTEM CHARACTERISTICS VARIED CONSIDERABLY ACROSS MEP SITES

As one might expect, the eight sites produced evaluation systems that look very different from each other when their operations are examined.

Table 8 shows where the sites fall on a number of descriptive dimensions. Clearly no typologies based on system operations emerge.

Individual case studies prepared on each site used the framework shown in Figure 4 to describe evaluation systems. Three sectors are distinguished: user of evaluation; subject of evaluation; and evaluation operations.

The evaluation operations of an MEP were sometimes funded by a number of sources, and sometimes carried out by different organization units. In some MEP systems, the Planning Agency's Evaluation Unit did not do the evaluation (e.g. Ventura), or they did it with outside assistance (e.g. Michigan's intensive evaluations). Other sections have already described the users and subject of evaluation (see the discussion of markets and Table 2). Here we will discuss the following characteristics of the evaluation system operations:

- evaluation system production levels;
- SPA/RPU Evaluation Unit product;
- standardization of the production process;
- evaluation methodologies;
- quality control mechanisms;
- user involvement mechanisms.

Only the last--user involvement in evaluation operations--was associated with success on the use criteria.

1. EVALUATION PRODUCTION LEVELS ACHIEVED

Table 9 shows the level of production associated with each site. Since the nature of the products varied enormously from site to site, one cannot compare the relative productivity of the systems. They produced different things for different purposes.

TABLE 8: MEP EVALUATION SYSTEM CHARACTERISTICS

	Primary User	Subject of Evaluation	Planning Agency Product	Standardization of Evaluation System	Staffing Arrangement	Cost/ Product	Study Method	Involvement of User in Production
Central Oklahoma	Subgrantees	SPA Projects	Study	Partial	In-house	Moderate	Descriptive	Low
Ventura	Local Officials	Criminal Justice System Activities	Technical Assistance	Ad Hoc	In-house	Low	Descriptive	High
Michigan (Standard)	Planning Agency	SPA Projects and Programs	Analysis of Reported Data and Study	Standard	In-house and Interns	High	Descriptive and Comparative	High
Michigan (Intensive)	Planning Agency	SPA Programs	Study	Partial	Contractor	High	Quasi-Experimental	Moderate
Illinois	Planning Agency	SPA Projects	Tabulations of Reported Data	Standard	In-house	Moderate	Tabulations	Low
Alameda	Local Officials	Criminal Justice System Activities	Studies	Ad Hoc	In-house	Moderate	Descriptive	High
Columbia	Subgrantee	SPA Projects	Technical Assistance	Ad Hoc	In-house	Low	Tabulations	High
Jacksonville	Local Officials	Criminal Justice System Activities	Study	Ad Hoc	In-house	Moderate	Descriptive	Moderate
Pennsylvania	Planning Agency	SPA Programs	Study	Partial	In-house	---	---	Low

FIGURE 4: EVALUATION SYSTEM OPERATIONS, INFORMATION FLOWS, AND RESOURCE FLOWS

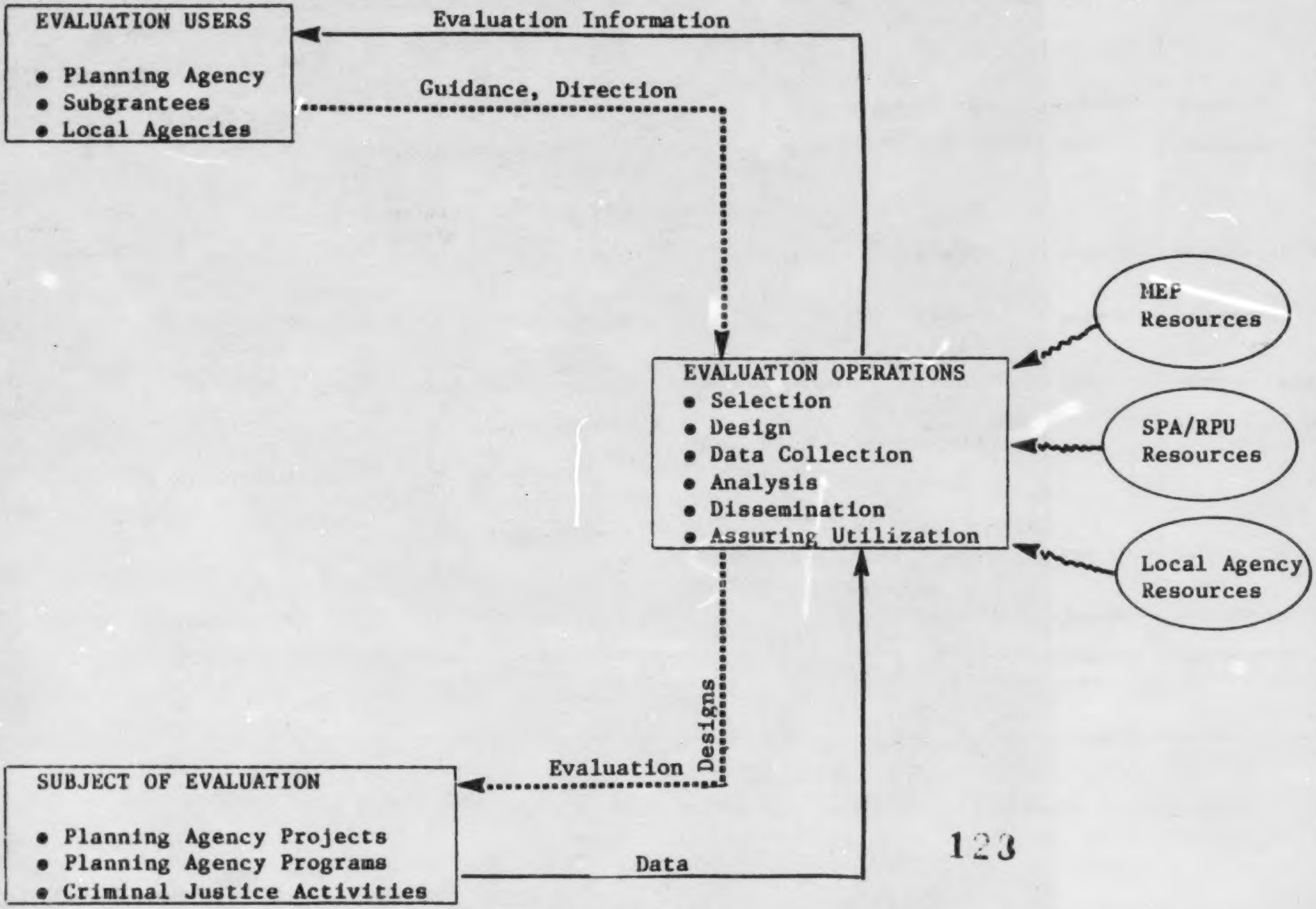

EVALUATION USERS
• Planning Agency
• Subgrantees
• Local Agencies
Evaluation Information
Guidance, Direction
MEP Resources
EVALUATION OPERATIONS
• Selection
• Design
• Data Collection
• Analysis
• Dissemination
• Assuring Utilization
SPA/RPU Resources
Local Agency Resources
Designs
Evaluation
SUBJECT OF EVALUATION
• Planning Agency Projects
• Planning Agency Programs
• Criminal Justice Activities
Data

TABLE 9: PRODUCTION LEVELS ACHIEVED BY MEP SITES

	MEP Budget*	Number of Analyses on Studies Completed	Were substantial non-MEP resources used to do the study?	Could future economies of scale be expected?
Central Oklahoma	$108,000	6	No	No
Illinois	$250,000	16	No	Yes
Ventura	$ 74,000	18	Yes	No
Alameda	$144,000	8	Yes	No
Michigan (Reporting System)	$142,000*	4	Yes	Some
Michigan (Intensive)	$190,000*	2	No	No
Columbia	$ 56,000*	4	No	No
Jacksonville	$ 85,000	7	No	No

---

*Two year grants include 50 percent second year match.

The site with the highest number of studies had one of the smallest grants, Ventura. High apparent productivity was achieved because Ventura played a technical assistance role; most of the costs were carried by local agencies.

The most expensive studies were the two intensive evaluations contracted out by Michigan at approximately $95,000/study. The simplest analyses were the tabulations of project process data produced by Illinois. The two reporting systems, Michigan and Illinois, have the possibility of achieving economies of scale if continued.

2. PRODUCT OF THE PLANNING AGENCY's EVALUATION UNIT

The MEP evaluation units were set up to deliver one of the following products:

(1) routine analyses of reported SPA subgrantee data--
  - Michigan and
  - Illinois;

(2) one-time studies of SPA projects or criminal justice system operations--
  - Michigan (Intensive),
  - Central Oklahoma,
  - Alameda, and
  - Jacksonville;

(3) technical assistance in evaluation to local agencies or subgrantees--
  - Columbia and
  - Ventura.

The two reporting systems produced different types of analysis. Illinois' products were largely tabulations of process data for individual projects. Michigan's produced project tabulations, plus an analysis of the program area.[1] The program analyses consisted of project comparisons using reported data and descriptive data collected on-site by the Evaluation Unit.

---

[1] These reports were partially completed, but not yet available as of February 1978.

The four sites that set up to routinely turn out one-time studies tailored these products to their own situation. The most sophisticated were the Michigan intensive evaluations which used quasi-experimental designs to estimate program effectiveness and impact (on crime). The other three sites turned out studies that were, for the most part, descriptive, with some use of before/after and planned/actual comparisons.

The technical assistance efforts of Columbia and Ventura were directed at two different audiences. Columbia was providing assistance to subgrantees on how to meet SPA evaluation requirements and make use of the resulting information. The technical assistance included: answering project questions about evaluation requirements, developing evaluation designs, collecting data and writing reports. The assistance was allocated across eight subgrants.

Ventura provided technical assistance to local agencies who were carrying out or wanted to carry out evaluations. In addition to the full-time evaluator, the MEP resources usually went to purchase the following types of service:

- computer time;
- consultant time;
- research assistant time;
- key punching, programming and
- travel.

The Ventura resource expenditure on a study ranged from a few days of professional time:

- Organizational Study of Police Department:

  --MEP Evaluator: three days,

--Consultant: one day ($135),
--Computer Time: $40;
--Key Punching: $10;

to the largest which required several man-months:

- Youth Services Project:

  --MEP Evaluator: six weeks;
  --Consultant: two to five days ($135 per day);
  --Students: (not available);
  --Office: $1200;
  --Supplies: $100;
  --Computer: $150;
  --Key Punching: $300.

The cost per product experienced by the sites can be classified as follows, using professional man-months as a measure of cost:

- Low - less than three man-months per product:

  --Ventura (technical assistance to eighteen studies),
  --Columbia (technical assistance to eight projects);

- Moderate - between three and twelve man-months per product:

  --Central Oklahoma (six project studies),
  --Illinois (sixteen analyses),
  --Alameda (eight studies),
  --Jacksonville (seven studies);

- High - greater than twelve man-months per product:

  --Michigan, Standard Evaluation (four program analyses),
  --Michigan, Intensive Evaluation (two program studies).

As noted earlier, the two reporting systems could reduce the cost/product by achieving economies of scale, while the other models could not.

3. STANDARDIZATION OF THE PRODUCTION PROCESS

The evaluation system has to carry out certain operations in order to produce information--selection, design, data collection, analysis, dissemination, and assuring utilization. The MEP systems differed on the degree to which formal methods and processes were developed to execute these operations. There were four kinds of prescribed structure given to a system:

- standard designs--specific instruments and analysis plans to be applied to projects of a given type;
- formal processes--major activities are layed out as a series of steps to be routinely executed;
- evaluation framework--typologies of measures, designs, etc. used to guide the evaluation work;
- none (ad hoc)--staff are made available to carry out the activities in a way which fits the situation.

Table 10 shows how prescribed the various systems were on different activities. The most prescribed systems were the two standard reporting systems--Illinois and Michigan--where most activities had a formal process, standard instruments, or a guiding framework.

Three systems were partially standardized--Central Oklahoma, Michigan (Intensive), and Pennsylvania. They used an evaluation framework (definitions, methods, etc.) or standard processes that were to be followed. Four of the systems were ad hoc--Ventura, Alameda, Columbia and Jacksonville. They made staff available to design and carry out studies required by the situation at hand.

4. EVALUATION METHODOLOGIES

The MEPs used fairly basic evaluation methods. There was no relationship between the type of design and utility. Descriptive studies were as useful as those using comparisons to make attribution or set norms.

For most of the sites--Central Oklahoma, Illinois, Alameda, Jacksonville, Columbia--the information produced was primarily descriptive, with some questions addressed through comparisons. The comparisons included planned

TABLE 10: METHODS FOR PERFORMING EVALUATION OPERATIONS IN THE MEP

Evaluation Operation	ACOG	Ventura	Michigan (Standard)	Michigan (Intensive)	Illinois	Alameda	Columbia	Jacksonville	Pennsylvania
Selection	Ad Hoc	Ad Hoc	SPA Selection Process	SPA Selection Process	(All Evaluable Projects Selected)	Ad Hoc	Ad Hoc	Ad Hoc	SPA Selection Process
Design	Framework to Guide Design	Ad Hoc	Framework and In-house Design Process	Contractor/ SPA Design Process	(Standard Process Was To Be Used during MEP)	Ad Hoc	Ad Hoc	Ad Hoc	Two Phased-Program and Evaluation Design Process
Data Collection	Types and Sources Specified by Framework	Ad Hoc	Standard Instruments and Project Reporting Requirements	Ad Hoc	Standard Instruments and Project Reporting Requirements	Ad Hoc	Ad Hoc	Ad Hoc	Ad Hoc
Data Processing and Analysis	Types Specified by Framework	Ad Hoc	Analysis Plans and Central SPA Data Processing	Ad Hoc	SPA Data Processing and Analysis System	Ad Hoc	Ad Hoc	Ad Hoc	Ad Hoc
Dissemination	Ad Hoc	Ad Hoc	Application Groups	Ad Hoc	Ad Hoc	Two Different Systems	Ad Hoc	Ad Hoc	Ad Hoc
Assuring Utilization	Ad Hoc	Agencies Request Help on Their Own Evaluations	Program Specialist Prescribed Functions	Ad Hoc	Ad Hoc	Responsibility Formally Delegated to User Group	Ad Hoc	Study Follow-up Process	Implementation of Optimum Models

versus actual, before and after, and comparison groups (clients, projects, etc.).

Ventura gives a good picture of the types of design employed since it was involved with 18 studies, most initiated by local agencies. Table 11 shows how many of the 18 used different type analyses.[1] The table shows that most of the measurement is of project process and outcome, and that most of the analysis involves no comparisons (descriptive), or are "planned versus actual" and "before/after." The impact analyses are usually analyses of crime data.

Michigan's intensive evaluation system used the most sophisticated designs. One such design is illustrated in Table 12. Strangely enough, in these studies, the most useful piece of information in the Michigan evaluation came from the measurement of process and comparison with grant applications (planned versus actual).[2] The contractor concluded that the two program areas studied were never implemented as planned. The apparent lack of project effects is due to implementation failure, rather than conceptual failure. In retrospect, the sophistication and cost of the evaluation design were not warranted, given the state of program implementation.

5. QUALITY CONTROL MECHANISMS

MEP sites did not use any standard validity or reliability checks on their data collection or analysis. Generally they relied on the process of using the data to identify inaccuracies and errors. This process involved returning processed data to those that had submitted them originally.

---

[1] The table was prepared by the MEP evaluator.

[2] This information raised issues about Michigan's administrative policies and practices, which led to changes in the role and time allocations of the SPA program staff.

TABLE 11: VENTURA MEP - FREQUENCY OF USE OF DIFFERENT DESIGNS (MEASURE AND COMPARISON) IN EIGHTEEN STUDIES

	Type of Measure*			
Type of Comparison	Input	Process	Outcome	Impact
Descriptive (No comparison)	7	9	8	4
Planned versus Actual	1	3	2	1
Before and After	2	6	7	6
Experimental Control Groups	0	0	0	0
Comparison Group	2	1	1	1

---

*Input: Resources going into a unit.

Process: Activities carried out by unit evaluated.

Outcome: Direct result of activity, e.g., change in client behavior.

Impact: Effect of unit on its environment, e.g., reduction in crime.

TABLE 12: SUMMARY OF MICHIGAN's EVALUATION DESIGN FOR YOUTH SERVICE BUREAUS

Question	Measures	Unit of Analysis	Data Source	Comparison
1. Do YSBs reduce crime and delinquency (impact on system)?	Reduction in reported offenses	Cities and counties	Uniform Crime Reports (UCRs)	YSB versus non-YSB jurisdictions
	Changes in juvenile justice system statistics--e.g., arrests	Cities and counties	Juvenile agency records	Pre-post YSB intervention
Do YSBs reduce crime and delinquency (in YSB clients)?	Reduction in recidivism	Client	Police and court records of clients	Client measures pre-post YSB
	Reduction in juvenile justice system contacts	Client	YSB records (MSU instruments)	YSB clients versus non-YSB comparison group
	Reduction in self-reported deviant behavior	Client	Self-reporting questionnaire (MSU instrument)	
2. Does diversion take place (system changes)?	Changes in juvenile justice system statistics--e.g., penetration	Cities and counties	Juvenile agency records	YSB versus non-YSB jurisdictions
	Changes in YSB statistics	Cities and counties	YSB records (MSU instruments)	Pre-post YSB intervention
	Changes in processing criteria and decisions of juvenile justice system agencies	Cities and counties	Juvenile agency records	

TABLE 12: SUMMARY OF MICHIGAN's EVALUATION DESIGN FOR YOUTH SERVICE BUREAUS (Cont'd.)

Question	Measures	Unit of Analysis	Data Source	Comparison
3. Does exposure to YSB prevent the development of the negative attitudes, self-concepts, etc. associated with the theory of secondary devance?	Attitude scales	Client	MSU instruments	Post-YSB arrestees versus pre-YSB arrestees
4. What do YSBs do, and why do they do it? (Was the project implemented as planned, and what factors contributed to the apparent success or failure of the project?)	Internal project characteristics (e.g., funding, objectives, structure . . .)	YSB	• YSB proposal and quarterly reports • YSB records (MSU instrument) • Structured interviews	Case history or case study type analysis Planned activity and results versus actual activity and results
	Immediate social context (e.g., juvenile justice system . . .)	YSB	Structured interviews (MSU instrument)	
	General social context (e.g., economic conditions . . .)	YSB	Structured interviews (MSU instrument)	
5. What kind of YSBs are most effective?	Project typologies: • organizational structure and program design; • community structure; and • inter-agency relationships.	YSB	Same as number 4	Contingency tables: • Project typologies are independent variables.
	Outcome criteria: • diversion; • client attitude; and • client behavior.		Same as numbers 1, 2 and 3	• Outcome criteria are dependent variables.

In both the Illinois and Michigan reporting systems, for example, the data tabulations were to be fed back to the subgrantees and given to the SPA program staff familiar with the project. Since the data were being used in project oversight, the expectation was that mistakes would lead to strange results which some party in the review and decision process would raise.

The only significant investment in a quality control function was made by Michigan on its standard reporting systems. Student interns were used to monitor the implementation of the system and the data flow. This included:

(1) On-site check of the appropriateness of the standard instruments to the project and the ability of the sites to collect the data.

(2) Explaining the instruments to each project.

(3) Visually checking the data submitted for obvious errors, incompleteness or inconsistencies.

(4) Discussions with projects over problems encountered.

Even here, however, there was no independent data collection to check the reliability and accuracy of the reported data.

6. MECHANISMS TO INVOLVE THE USER

The degree to which users were involved in the evaluation effort can be described as reactive, active direction, and active participation:

- Reactive: User may be an information source during design work, and may review designs developed by the evaluators, but does not provide guidance on substance[1] of the design:

---

[1] Substance refers to research questions and the measures and comparisons proposed to answer them.

(1) Central Oklahoma;
(2) Illinois;
(3) Columbia.

- Active Direction: Users gave guidance and assistance on the substance of the design and analysis:

  (1) Ventura;
  (2) Michigan (Intensive);
  (3) Alameda (Pre-trial Studies).

- Active Participation: Users carry out one or more of the evaluation operations, such as design, data collection, analysis:

  (1) Ventura;
  (2) Michigan (Standard);
  (3) Alameda (Juvenile Studies).

In Chapter II we described how the "active direction" and "active participation" led to user ownership of the study and contributed to the uses.

Ventura used both active participation and active direction: (1) let the user initiate the study and carry out parts of it (active participation); and (2) develop a strong interpersonal working relationship with the user (active direction). The success of the latter was brought out by interviews with users. They described the evaluator as follows:

> He enjoys a considerable amount of rapport from organizational studies (of police departments) he did before MEP. When he talks, police chiefs listen. He has credibility and trustworthiness among police chiefs, etc., which grew out of work he did for them.
>
> He has credibility among the various criminal justice elements and can anticipate their data needs . . . and he's got the credentials. He is a likeable guy whose credentials and personality enable him to get along with the various elements in the criminal justice system. (The good things about him are) his professional qualifications, his energy, and personality.
>
> He could relate to some of the concerns of the operational types. He knows a lot about a lot of different things. Operational managers could

> see they were getting something out of it. They could trust him, and knew the study would not be used against them. He could really help them and not be a nuisance--this started based on his credibility as an individual.
>
> In Witness Utilization, he sat down and listened to what everybody wanted and what our concerns were. He gets along well with people--he was objective--he understood the system reasonably well and he related well.

Michigan's Standard Evaluation System used a combination of staff training and work groups to involve the users, i.e., the SPA program specialists in "active participation." All SPA staff received training in evaluation (80 man-days of staff time) as part of the MEP, and all staff participated in work groups (75 man-days of staff time) that produced the standard instruments. Subsequent work groups (post MEP) were restricted to program staff with responsibility for the program area and evaluation staff. The data collection and reporting instruments are designed by the program specialists in the work group. Evaluation staff operate in a supportive role.

The Michigan Intensive Evaluation System used SPA work groups (25 man-days of staff time) to give "active direction" to the contractor during the evaluation design phase (3 to 6 months, 20 percent of contract costs). The group interaction was augmented by surveys of management and group specialists carried out by the contractor. (The surveys were judged not as useful as expected by the contractor and MEP staff.) The work groups consisted of the program specialists for the area under study, MEP staff, contractor and, on some occasions, top management. The work groups did not participate in the actual design work. The design was more the contractor's, but the SPA staff for the most part felt it would meet their needs, and reflected their discussions with the contractor.

Alameda also used two different mechanisms to involve the user. Both gave the user administrative responsibility for the study team and its product. In one case (Adult Pre-trial), the studies were done for existing government organizations with statutory responsibilities and authorities. These bodies gave "active direction" to the study, through all phases of the work.

In the second case (Juvenile), a committee was established to direct the study. Representatives from local agencies, advocate groups, and the public at large were brought together. The committee met regularly--usually monthly--to review the progress of the study and, often, to redirect certain activities, and add or delete questions to be addressed. The committee reviewed all products, sponsored public hearings on the reports and, after review and debate, wrote conclusions and recommendations. They ended up becoming "active participants."

The reactive approach requires little involvement of the user, and no commitment from the user on what information is really needed. Typically, following an interview with a potential user and review of documents, the evaluator will propose a design for comment or approval. Under this process, if the design is accepted, you cannot be sure that the user really wants it or just is not interested. Perhaps anything would have been accepted. The MEP experience shows that the reactive mechanism usually finds the evaluator guessing wrong, and the identified user not really interested.

E. BRIEF DESCRIPTIONS OF THE EVALUATION SYSTEMS

In the previous section we contrasted the MEP evaluation systems on a number of characteristics. The following are brief descriptions of each evaluation system, stressing its major characteristics.

1. A STANDARDIZED SPA PROJECT REPORTING SYSTEM DEVELOPED BY THE EVALUATION UNIT: ILLINOIS

The Illinois system produced tabulations of data reported by subgrantees using standard instruments. The instruments had been produced for a variety of project types during the MEP.

The operation of the system primarily involves data processing. The system does not have staff invested in either implementation of the reporting system on site, or maintaining quality control over the data flow. Data processing was to be automated, and the analysis was to utilize socioeconomic data bases; however, these activities were never routinized. The standard instruments appear to cover half of the funded projects, and data reporting by subgrantees was a problem.

The Illinois reporting system can be viewed as a "minimum" cost version of an evaluation system frequently attempted by SPAs.[1]

2. A STANDARD SPA PROJECT REPORTING SYSTEM DEVELOPED BY PROGRAM STAFF: MICHIGAN

The Michigan standard evaluation system is similar in principle to Illinois, but different in three areas. First, the Michigan system relied heavily on interns to develop and implement the system. The interns went on site to introduce the data to the projects and exercised quality control over the data flow. Second, the interns collected descriptive data on site and used these data, along with the reported data, to produce reports. Third, the state program staff designed the instruments, helped introduce the instruments to the site, and participated in the analysis.

---

[1] For other examples, see Monitoring for Criminal Justice Planning Agencies, John Waller et al., The Urban Institute, August 1974.

Design of instruments is an on-going function of the Michigan system; it was not for Illinois. The Michigan system, as tested, can be viewed as a "high" cost version of a reporting system model.

3. PHASED CONTRACTS TO DO INTENSIVE EVALUATIONS: MICHIGAN

Michigan's second evaluation system attempted to estimate the effects and impact of programs by investing a large amount of resources in data collection and analysis. A contractor was hired to test the concept.

The contractor had a two-phase contract. The first part was a design phase in which he was to analyze the agencies' information needs and the nature of the projects in place. The design phase took 20 percent of the contract resources. Michigan also set up management and program staff groups to interact with and review the work of the contractor.

4. IN-HOUSE STAFF DO SHORT-TERM EVALUATION ON DEMAND: JACKSONVILLE

The MEP in Jacksonville was designed to expand an existing evaluation system to provide additional information to criminal justice decision makers who, over a several year period prior to the MEP, had learned to rely on the RPU for planning information. The existing system used a two person staff to conduct short (six weeks) evaluations of LEAA funded projects.

The RPU planned to expand the existing system by increasing the size of the staff, using consultants to advise and assist the staff, changing the unit to include non-LEAA funded activities, and increasing the number of evaluation studies conducted. During the MEP, the evaluation unit of the OCJP tripled in size, consultants were used for only a short while, and the unit of analysis was changed. However, the number of studies conducted was the same as in the year prior to the MEP. The studies tended to be descriptive studies of existing criminal justice activities that provided information for local agencies to use in altering the activities or planning for future activities.

5. IN-HOUSE STAFF DO PROJECT EVALUATIONS USING STANDARD FRAMEWORK: CENTRAL OKLAHOMA

The purpose of Central Oklahoma's MEP grant was to develop, field test and package for use by others a method of evaluation that would be useful for local government agencies. The work was performed by an in-house staff with consultant support.

With the MEP funds, Central Oklahoma assembled a staff and carried through with its plan to produce an evaluation method. The method consists of a framework for defining objectives, measures and analyses, and guidelines for using the framework to develop project specific evaluation designs. During the MEP grant period, the method was used to guide the performance of six project-level evaluations by the ACOG staff. Separate reports were prepared for each evaluation documenting the results obtained and the method as applied to a specific project. The field test experiences were used to develop a training manual to assist local agencies in conducting evaluations. Training sessions were to be given, but MEP grant time and resources were exhausted before training could be provided.

There is an organizational difference, however, between the evaluation method as implemented during the field test, and as originally envisioned or described in the training manual. The field test involved the evaluations being conducted by Central Oklahoma personnel; whereas the documentation produced on the method indicates that local agencies would conduct the evaluations of the projects they were operating. Also, in the documentation produced on the method, it is assumed that the evaluation design and project design develop concurrently, and that special data recording procedures could be incorporated into the project. These conditions were not feasible during the MEP field test of the evaluation method. Due to lack of data and appropriate project activities, several types of analysis specified in the framework were not feasible.

6. RPU RESEARCH TEAMS DO STUDIES UNDER THE DIRECTION OF USERS: ALAMEDA

With the Model Evaluation Project grant, the Alameda RPU demonstrated the feasibility and value of evaluations of a particular type for local criminal justice agencies. As initially envisioned, the evaluations would involve:

- an assessment of the impact of selected innovative projects on a given problem in relation to other projects with similar goals, and to the more traditional institutionalized approaches to the problem, and

- the development of a cost effectiveness analysis of selected innovative projects in comparison to the cost effectiveness of more traditional institutionalized programs for addressing similar problems.

The studies were performed on an in-house basis under the direction of the RPU research staff, with extensive interaction among components of the local criminal justice system.

A series of studies were conducted in two areas of the criminal justice system: juvenile justice and adult pre-trial services and detention. Existing organizations and committees were identified or created, and used as customers and sources of guidance for the studies. This involvement of users in the evaluation activity is one of the unique aspects of the Alameda Project. Other unique aspects were the scope of the studies undertaken, and the relationship of the studies to controversial local issues. These aspects resulted in a level of interest in, and reaction to, the studies significantly higher than for many other MEP grantees.

Several of the planned studies were completed during the MEP grant period and used in subsequent policy and management deliberations and decisions by local criminal justice agencies. Portions of the planned analyses could not be completed, however, with available resources or within the projected time.

7. ADVICE AND ASSISTANCE ON EVALUATION TO SUBGRANTEES: COLUMBIA

The Columbia MEP Project consisted of a regional resource person providing advice and assistance to local officials seeking or evaluating LEAA grants. Initially the plan was to develop evaluation designs that would be implemented by local officials with assistance from the RPU evaluator, and which would produce information useful to local officials.

The aim was to develop and institutionalize local appreciation and use of evaluation and the capability to perform evaluations. However, the State Planning Agency has specific evaluation requirements and designs that are placed on most grantees. When the regional evaluator was unsuccessful in obtaining the cooperation of local officials in expanding from the SPA minimum requirements, a posture of facillating local complaince with the state mandates was adopted. The facillation involved working with grantees to: negotiate acceptable implementation plans; develop procedures for obtaining and analyzing data; review products; and provide some direct assistance by collecting and analyzing a limited amount of data. All products except one were project reports to the SPA. The exception was a report to the SPA jointly authored by project staff members and the regional evaluator.

8. EVALUATION TECHNICAL ASSISTANCE TO LOCAL AGENCIES ON REQUEST: VENTURA

The Ventura MEP Project primarily consisted of one full-time resource person. The regional resource person had available some funds that were used selectively: to support a few consultants and part analysts used periodically; operate a short lived seminar series; and sponsor a small amount of travel and course work for local personnel. The resource person became involved in 18 evaluations by performing a variety of tasks. A few evaluations he handled entirely on his own, but for the majority his role was to perform or assist in the performance of some key steps in the evaluation process. Examples of the selective involvement are: develop an evaluation design that others implement; or analyze the data collected by others.

F. CONDITIONS WHICH MAY AFFECT IMPLEMENTATION

Only one MEP site failed to implement an evaluation system. Three met their original expectations, and four changed their expectations but were able to implement some version of an evaluation system. Given the success of MEP sites at developing evaluation capability, one can conclude that simply earmarking funds for evaluation is very likely to result in evaluation systems and in evaluation information. Most sites made adjustments as problems arose, and carried through development and implementation efforts.

Contrasting the site that failed with the other seven provides a few lessons about conditions that might affect implementation:

- The SPA management support can make a difference to the successful implementation of a system.
- Making a commitment to a real user may drive one to produce evaluation.

Management's role can be illustrated by contrasting Michigan and Pennsylvania. The management of the Michigan SPA made it very clear that it supported the MEP, and expected staff to participate. When staff attendance at training sessions began to fall off, management spoke out and reversed the trend. The Pennsylvania SPA, on the other hand, caused or allowed the staff of the MEP to be drawn away to non-evaluation activities. The staff was asked to carry out SPA administrative functions. This diffusion of MEP resources and initiatives helped disable the effort in Pennsylvania.

Pennsylvania was also the only site where a real user--i.e. people carrying out functions--was never identified. All other sites made some commitment at some point to provide assistance or information to some person. This never occurred in Pennsylvania because the new SPA decision processes the MEP expected to be created were not established during the two year grant period.

IV. PROBLEMS TO BE EXPECTED IN IMPLEMENTING AND OPERATING EVALUATION SYSTEMS

The MEP sites ran into a number of problems that threatened implementation of the evaluation system. Sometimes the problems were disabling and required a change in strategy and expectations. In the end, most sites were able to resolve or adjust to these problems and test an evaluation system.

This chapter describes the most significant problems encountered by one or more of the MEP grantees. These problems are those that a good evaluation manager should expect to occur, and should be prepared to resolve. There are no general solutions; evaluation management will have to exercise judgment on how best to resolve the problems when encountered. The problems are listed in Table 13, and each is discussed below. At least one example is given of how each occurred.

A. STAFF TURNOVER IN THE EVALUATION UNIT WILL CAUSE DELAYS

Staff turnover was significant in four of the eight completed MEP sites. Four people passed through Michigan's evaluation manager position, and it was vacant for several of the twenty-four months of the grant. Columbia, a one position MEP, had two people pass through the position. Nine people passed through four MEP funded slots in Jacksonville. The

TABLE 13: OPERATIONAL PROBLEMS THAT EVALUATION MANAGEMENT SHOULD BE PREPARED TO RESOLVE

A. Staff Turnover in the Evaluation Unit Will Cause Delays and Disruptions

B. Civil Service Rules May Cause Delays

C. Projects May Be Reluctant to Report Data Needed for Evaluation

D. Project Files May Contain Less Data than Expected or Required for the Planned Evaluation

E. Computer Processing of Data Can Be a Time Consuming, Error Plagued Undertaking

F. Contracting out Evaluations Has Its Own Set of Problems

G. The Identified User Can Add to the Cost of Evaluation by Having Unclear or Shifting Information Requirements

H. Diversion of Evaluation Unit Staff to Other Planning Agency Functions Is a Danger Signal

MEP manager in Pennsylvania left and the position was vacant for almost two years.

Staff turnover blunts the momentum of a capability building effort and, when it is in a top management position, it can stop progress. Pennsylvania's Evaluation Unit identified the absence of an MEP manager as a major problem:

> There was no Chief of Program Evaluation for most of the grant period. Without that slot filled, there was no one available with sufficient time to guide the effort and protect the staff from outside requests to do non-MEP tasks.[1]

Michigan, on the other hand, limited the negative effect of a high turnover in a similar position when the Director of Planning assumed some of the MEP manager's responsibilities. Michigan made sufficient adjustments to allow implementation to go forward close to schedule.

Evaluation managers should expect some level of turnover among in-house and contractor staff. The question is whether turnover will seriously affect the evaluation system. Contingency plans can be developed to deal with turnover in key positions.

B. CIVIL SERVICE RULES MAY CAUSE DELAYS

Civil Service regulations introduced several problems—bringing recruits through the procedures took longer than expected, sometimes the type of staff required could not be obtained under Civil Service rules,

---

1. John W. Scanlon and John D. Waller. Final Report on the Model Evaluation Program of the Pennsylvania Governor's Justice Commission," Paper 5091-2, Washington, D.C., The Urban Institute, April 1978.

and in one case several vacant staff positions crucial to the MEP fell under a no hiring freeze.

Several of the MEP sites experienced start-up problems because of Civil Service procedures. They ended up starting late, or at a much slower rate than anticipated. In one agency people were put into vacancies for brief periods to conform to Civil Service rules, although they were not really qualified or interested in the MEP.

Pennsylvania was most seriously affected by Civil Service. The MEP was effectively at half strength (five positions authorized) for eighteen of the twenty-four months. One reason was a State hiring freeze. Near the end of 1975 and to the second half of 1976, the SPA was under a State hiring freeze. When the hiring freeze was lifted, new Pennsylvania Civil Service procedures slowed hiring and disrupted operations. Both vacant positions and filled positions were affected. Onboard staff had to be reexamined under the new procedures, and for a while, their positions and ratings were in jeopardy.

The manager of an evaluation system should take care to understand the Civil Service system under which he or she operates, and know the constraints it will place on implementation.

C. PROJECTS MAY BE RELUCTANT TO REPORT DATA

The two reporting systems tested had very different experiences in establishing a flow of data from subgrantees.

Illinois was not able to get all projects to send in data. In some cases, subgrantees were told that the data would not be used in refunding decisions if they sent them in. Illinois had not anticipated the refusal of subgrantees to report, and did not have a strategy to deal with it.

Michigan had good reporting from all projects in the test. Reporting requirements were written into subgrants, and the MEP Unit invested time and effort in introducing the projects to the reporting instruments.

Since timely and accurate reporting is crucial to this type of evaluation system, the evaluation manager should make sure at the start he or she has the authority, resources and leverage required to get adequate reporting.

D. PROJECT FILES MAY CONTAIN LESS DATA THAN EXPECTED OR REQUIRED FOR PLANNED EVALUATION

Central Oklahoma planned to carry out evaluations of projects relying on existing data sources--primarily data routinely collected by the projects. The project data bases turned out to be insufficient for implementing the designs that had been developed.

Central Oklahoma had developed an approach to project evaluation which:

- viewed every project as having three types of objective: immediate, intermediate, and ultimate;
- provided a set of five criteria that would be used to measure project performance relative to its objectives: effort, level of performance, adequacy of performance, efficiency, process analysis; and
- used routinely collected project data.

The limitations of the existing data prevented the MEP staff from conducting about half the analyses intended. Several of the evaluation criteria could not be applied in the analysis. Project performance on the immediate objectives was usually estimated; however, assessments of performance could not be made for about one-third of the intermediate and ultimate objectives. The process and efficiency analyses could not be performed due to the lack of data and appropriate comparisons.

Michigan's intensive evaluation contractor also found project and agency data bases weak. For certain variables, the contractor was forced to set up his own measurement system in projects.

The evaluation manager should quickly and early on size up the actual data bases that will be available. An approach to evaluation or an evaluation design should not be settled on until the feasibility and cost of getting the data is well known and verified.

E. COMPUTER PROCESSING OF DATA CAN BE A TIME CONSUMING, ERROR-PLAGUED UNDERTAKING

Both MEP sites that relied heavily on computer processing, Michigan and Illinois, ran into long, unanticipated delays due to use of the computer.

In the Illinois case, the data processing was handled in-house. The Evaluation Unit was not prepared to process the data that were submitted. Backlogs quickly developed, and several analyses were delayed when data had to be reformatted before processing. The backlog became a problem because it prevented the quick return of processed data to projects for verification and use. This eliminated the primary quality control mechanism, reduced the initiative of projects to submit data, and prevented one of the anticipated uses from occurring. The size of the backlog was also a problem, since it grew to the point were resources did not exist to reduce it and use all the data that had been collected. Earlier realization of the processing problem could have been used to support decisions on reducing the data collection requirements, and prevented the expenditure of resources on data that were not being processed. Once the step was taken to eliminate most data collection requirements, the Evaluation Unit made a commitment to process much of the data that had been collected. The extent to which the commitment will be realized is unknown.

Michigan contracted with a University computer center to process its data. MEP student interns worked with a center programmer to program and run the data. This phase of the evaluation system introduced several months of delay and ran into a number of difficulties. The interns described the problems as follows:

- The interns had little experience in programming. The job was given to them with no supervision. They had a difficult time doing the work.
- A computer center programmer would work out bugs with the interns. This interaction took a lot of time.
- The computer center erased a tape and misplaced others. The center was moving its hardware during this time.

The delays in the processing of data had negative effects on both MEP sites. It affected the reporting rates from projects, and it disabled attempts to verify (quality control) the data through a review process. In Michigan, where reporting by projects was generally good, some projects stopped sending in data for a period because the SPA was not living up to its agreement to provide feedback. The unanticipated delays threatened the agreements and cooperation upon which a reporting system must be based.

Many of the problems encountered at the two sites might have been avoided by careful up-front planning and more aggressive management. Upfront planning could have established more realistic time schedules to use in making feedback commitments to subgrantees and SPA users.

F. CONTRACTING OUT EVALUATIONS HAS ITS OWN SET OF PROBLEMS

One solution to a number of operational problems is to contract out the evaluation effort. For example, a contract is one way to avoid problems raised by Civil Service procedures, and it is a mechanism for quickly acquiring a technical capability that could not be established in-house. However, contracting has its own set of problems, and does not resolve many of the operational problems faced by the evaluation system.

Four of the twelve MEP sites relied significantly on contractors: Michigan for its intensive evaluation; Virginia and New Hampshire for developing monitoring and evaluation systems utilizing standardized data collection instruments; and Washington for provision of the training and technical assistance. All three ran into some difficulties: Michigan resolved its problems; Virginia terminated the contract; New Hampshire and Washington, after several contracting delays, were still fieldtesting the system developed when this document was prepared.

Michigan's contractor ran into staffing problems on its senior positions. This was one factor contributing to a three to four month extension of what had been planned as a three month design phase. Once the contractor presented an acceptable staffing plan, Michigan approved the evaluation designs and work proceeded. Michigan had set up a mechanism which allowed the contractor and SPA personnel to interact as the evaluation design evolved. Thus, when time came to make a decision to go ahead, the Evaluation Unit knew what it was buying. Moreover, prior to the final decision to go ahead, the evaluation coordinator spent a day with the contractor going over the design and data collection carried out to that point. Final approval rested on the contractor and Michigan agreeing to a staffing plan.

Virginia's contractor was to perform most of the work for developing the monitoring system--data collection and analysis procedures. The contractor began by conducting a needs and capability assessment. The SPA staff disagreed with the preliminary results of the assessment, and the effort was abandoned. Work then focused on the development of measures and data processing procedures. Again initial products did not coincide with SPA expectations, and a decision was made to terminate the contract.

An Evaluation Unit should be prepared to invest a significant portion of its staff time in working with contractors. Hiring a contractor creates a need of communication and decision mechanisms between the Evaluation Unit and contractor. Michigan had such mechanisms and was able to work effectively with its contractor.

G. THE IDENTIFIED USER CAN ADD TO THE COST OF EVALUATION BY HAVING UNCLEAR OR SHIFTING INFORMATION REQUIREMENTS

Involving the user in the evaluation process increases your chances of producing useful information. It may also increase the cost of doing evaluation. The experiences of Michigan's intensive evaluation and Alameda's juvenile studies are instructive.

Michigan's intensive evaluation spent about 20 percent of the contract resources on the design of the studies (about $40,000 out of $200,000 for two studies). The contractor found it very difficult to get the State Planning Agency to identify its information needs. The contractor described the problem as follows:

> We first tried to get a feel for Michigan's decision-making needs and issues. I'd say we failed. We never resolved the great diversity of opinion in the Michigan SPA. The diversity of opinion was on what Youth Service Bureaus were, and what to look at.
>
> The message I got was that the Office of Criminal Justice Planning wanted to look at a variety of things, not sink its dollars into one piece, for example, just a study of secondary deviance. So we developed a 'model' to address various needs and interests. It showed the relationship among the various effects that people were talking about. Spreading our resources across a lot of questions means we can be less definitive on any one.

The contractor work had to produce a design that reflected the diversity of views in the SPA and, judging from the user interviews, was successful. The point is that the task was more difficult and expensive than originally expected by the contractor or the MEP staff.

In Alameda the RPU management and study team identified as a serious problem the tendency of user groups to broaden the scope of the study efforts. Many members of the Juvenile Justice Committee set up to direct the study were not familiar with the activities being studied, and had to be "educated" by the study staff. The Committee also identified new questions to be explored and additional data sources to be utilized as the study proceeded. This tended to divert the staff from the original plan and delay schedules. The studies became so broad that several components could not be completed on time.

An evaluation manager should budget for involving the user. The process of interaction, education, and consensus building is always time consuming. A significant amount of Evaluation Unit management time and staff time should be set aside for such activities.

H. DIVERSION OF EVALUATION UNIT STAFF TO OTHER PLANNING AGENCY FUNCTIONS IS A DANGER SIGNAL

The sites that were most successful in implementing evaluation systems --Michigan, Ventura, Alameda, Central Oklahoma, Jacksonville--kept the MEP Evaluation Unit resources from being absorbed by other planning agency functions. Several MEP sites did not do this. They allowed staff to be diverted to other work, and by doing so, disabled or delayed the development of the evaluation systems. Virginia and Pennsylvania are examples.

With the MEP, Virginia's Evaluation Unit went from a staff of one to a staff of four. The new and existing evaluation staff were integrated into a single unit, with all members involved in a range of activities. Some activities were associated with developing the new MEP system, while others involved the staff in grant application reviews and short-term evaluations for the SPA. The tendency to treat the MEP staff as a regular part of the SPA staff resulted in a variety of demands being placed on them, and diverting them from MEP tasks on several occasions. The diversions are cited by the MEP staff as one reason that delays have occurred in the development effort.

A significant part of Pennsylvania's available MEP manpower went into supporting State office planning and administration functions. During the two year period, MEP staff were involved in:

- preparing parts of the Comprehensive Plans;
- developing funding guidelines for the program categories in the 1977 Comprehensive Plan;
- reviewing subgrant applications;
- dealing with several project evaluations for which the SPA was responsible; and
- serving as trouble-shooter, advisor or liaison on issues that came up related to evaluation or information systems.

These efforts were undertaken on an ad hoc basis--the situation would arise and the Evaluation Unit would respond. Staff involvement in these activities was always rationalized as contributing the experience and information necessary to do program evaluation. However, in most cases, the efforts were eventually judged to be non-productive for MEP purposes, and involvement would be phased down. Pennsylvania's inability to set up any evaluation system is closely linked to the way it dissipated its MEP resources on unrelated activity.

Clearly, the evaluation manager should view requests and assignments to help out on other Planning Agency functions as a danger signal. Developing, implementing and operating an evaluation system is a full-time job in itself. Diversion of staff is a sure way to stop its progress.

I. NO STANDARD PROBLEM AND NO EASY ANSWERS DEMONSTRATED

The above problems and the lessons summarized in Chapter I are only a synthesis of those encountered during the MEP and recurring daily in most agencies. Each problem was encountered in a particular context and required solutions tailored to the situation. The MEP experiences demonstrated that it is feasible to achieve an evaluation system which produces useful information, but there are no easy ways to solve problems that will be encountered. Management has to want the evaluation results, and be willing to invest time and resources in obtaining information if evaluation is to have utility. Evaluation cannot make a difference by just existing. Users and uses identified prior to initiation of evaluation studies enhance probability of utility.

The MEP experiences are directly applicable to those agencies that operate on annual or work plan or funding bases. All MEP sites were limited to a one or two year grant. Under those conditions, delays or other problems encountered prevented the development of planned products prior to grant termination. Since most sites were not continued, this situation resulted in some delayed products never being produced. Agencies that do not have production or funding deadlines may not encounter this situation. However, the problems just discussed and the lessons presented in Chapter I will affect the productivity of most evaluation systems to some extent and, consequently, are considered applicable to most government agencies.

☆U.S. GOVERNMENT PRINTING OFFICE: 1979-281-380/1611